This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.
Conceptualising A New Theoretical Framework of Inclusion/Exclusion and Gender-Based Violence Within Higher Education in the United Kingdom: A Mixed-Methods Approach Embracing Computational Social Science

ELLEN FRANK DELGADO

DOCTOR OF PHILOSOPHY - SOCIOLOGY
SCHOOL OF SOCIAL AND POLITICAL SCIENCE
THE UNIVERSITY OF EDINBURGH
2023
Declaration

I declare that this thesis has been composed by myself and that the work is my own, except where explicitly stated in the references or the acknowledgements. This work has not been submitted for any other degree or professional qualification except for the one specified.

Ellen Frank Delgado

15 October 2023
Abstract

This mixed-methods computational research concerns itself with harmful social phenomena that are difficult to measure; specifically, gender-based violence and inclusion/exclusion within one higher education institution in the United Kingdom. The thesis situates both social phenomena as parts of inequality regimes, as defined by Acker (2006). Tomaskovic-Devey & Avent-Holt (2019) emphasise how inequality regimes often employ covert mechanisms to maintain unequal social hierarchies. This work proposes a new theoretical framework to highlight this subtlety. Furthermore, the framework described is an attempt to overcome researchers’ overreliance on individual perception when exploring these phenomena by placing more emphasis on group-level inclusion/exclusion and gender-based violence. The framework also provides a foundation for an exploration of organisational interventions and puts forth an opportunity to innovate methodologically by borrowing from the field of data science. Utilising computational social science in a mixed-methods inquiry, the thesis investigates changes in perceptions and behaviours related to the implementation of one gender-based violence training intervention and two inclusion training interventions. These types of interventions are increasingly popular despite limitations in the methods used to evaluate them. As more higher education institutions employ such interventions, it is crucial to understand how the interventions may affect their unique and covert inequality regimes in place. The research uses these innovative methods in a proof of concept to highlight that gender-based violence and exclusion will continue to persist without an acknowledgement of the larger social context of academia in the United Kingdom. In this way, it becomes clear that higher education institutions must attempt to reconcile neoliberal forces such as performativity, growing preoccupation with financial gain, and other neoliberal logic in order to efficiently mitigate inequality regimes and render such trainings more effective.
Lay Summary

This research is a mixed-methods computational investigation into inclusion/exclusion and gender-based violence within a higher education institution. I explore the effects of two types of training interventions that are increasingly popular within academia and beyond: inclusion training, as well as sexual consent and active bystander training. These two types of interventions are the university’s strategic attempt at mitigating exclusion and gender-based violence alike.

In this thesis, I use data that is both qualitative and quantitative. Working mostly with first-year student participants, the research involved surveys, focus groups, individual interviews, and observations. However, at the crux of this work is an emphasis on research that does not over-rely on individual perceptions. Therefore, I propose my own theoretical framework that subsequently looks at group-level covert behaviours and perceptions, in addition to individual perceptions. To analyse the data, I employ thematic analysis, statistical analysis, and computational social science. More specifically, I advocate throughout this dissertation for the incorporation of computational text analysis and social network analysis into inclusion/exclusion and gender-based violence research.

The findings suggest evidence that the trainings cannot be the only solutions to combatting an organisation’s harmful status quo. With that, the trainings address some inequalities, but there are stark remaining inequalities. As a result, this dissertation stays rooted in practical steps forward for the higher education institution and higher education as a whole, keeping in mind the larger social context of UK academia. In particular, my work presents a way forward to further address the nuances of inclusion/exclusion and gender-based violence with more computational research. At a high-level, I also urge higher education to dismiss neoliberal pressures that promote performativity, foster market-driven logic, and ultimately, allow inequalities to persist covertly.
Acknowledgements

As much as my name is on the cover page, there are so many other people who supported me on this journey. I would like to extend my gratitude to my supervisors, Dr. Tod Van Gunten, Dr. John Amis, and Dr. Alison Koslowski. Their feedback and direction were instrumental in helping guide me through the last three and a half years. Much gratitude is owed to Tod in particular who believed in me and my work from the very first email I sent to him. I also wish to thank Dr. Chris Barrie and Dr. Samer Abdelnour who led my first year board examination and offered an initial round of feedback and encouragement. Other academics who have offered their advice, sent me journal articles, listened to my ramblings over (mostly virtual) coffee, and offered their support in countless other ways do not go unnoticed: Dr. Cindy Pace, Dr. Beverly Tarulli, Dr. Grant Jarvie, Dr. Paul Widdop, Dr. Richard Brodie, Dr. Seongsook Choi, Dr. Bernardo Ferdman, Dr. Jasmien Khattab, Dr. Hans van Dijk, Dr. Sameer Srivastava, Dr. Shamus Khan, and Dr. Courtney Cogburn. A special thank you also must go to those who helped me when R Studio routinely decided to hate me: Dr. Ugur Ozdemir, Dr. Eric Knudsen, Jake Barrett, and of course, Daan Paardekooper. My participants, many of whom are first year budding academics, were also my greatest source of knowledge and inspiration.

Throughout this process, my friends (so many on their own PhD journeys) have continued to be my pillars. You each know who you are! A special thanks goes to Isadora Dullaert and Elif Buse Doyuran, who have been sociology angels from Day 1 of our Zoom-doomed cohort. I also wish to thank Lauren Davis who has been one of my biggest supports since we spent our Saturdays in 2019 in overpriced cafés working on PhD applications together. Her ability to be forever supportive even at a distance makes her friendship one of my most cherished. I especially want to thank Madison Gerdes who makes me laugh even when our work feels overwhelmingly depressing and like our PhDs will never end. I also owe so much to Jo Becker who inspires and motivates me daily through her balancing of research and activism. Many academics could learn from her as I’m grateful to do. Charlie Wackwitz, I feel so lucky for the stars aligning and bringing us both serendipitously together in Edinburgh. His laughter and infectiously positive attitude are simply the best. I also genuinely think the universe knew I needed a friend in Emma McDowall because I still feel so lucky to have met her my first day walking around Edinburgh. Rarely as a grown adult do I randomly meet a stranger, end up chatting with them for two hours (*illegal at the time- sorry, Nicola Sturgeon*), and have them become one of the first people I talk to about everything and anything. Thank you to each of you.
This thesis would simply not have been possible without my parents. Pursuing an unfunded degree while trying to make ends meet during a cost-of-living crisis does not go without its challenges. Beyond their financial support, I thank them for instilling in me a value of curiosity and education from a young age.

To my partner Tijsje, I am truly so grateful for you. The last 15 years of friendship turned partnership have been a highlight of my life. This research and thesis likely would have been entirely differently shaped without your unwavering open ears. Thank you for being understanding, always.

Finally, it should go without saying, but I would vehemently not recommend to a friend, nor even an enemy, to begin a PhD during a global pandemic and national lockdown. If history ever repeats itself though, I highly recommend 1) surrounding yourself with good supportive people and 2) getting a dog. With that, to my dog Lottie: I sincerely hope this degree helps me put the highest grade top shelf organic non-GMO dog food in your dog bowl- you deserve it.
# Table of Contents

Declaration ................................................................. 1  
Abstract ........................................................................ 2  
Lay Summary .................................................................... 3  
Acknowledgements .......................................................... 4  
Table of Contents ............................................................. 6  
List of Figures and Tables ..................................................... 10  

Chapter 1: Introduction ......................................................... 14  
  1.1. Research Aims and Questions ........................................ 19  
  1.2. Positionality and Reflexivity .......................................... 20  
  1.3. Structure of Thesis ..................................................... 22  

Chapter 2: Theoretical Overview ............................................ 25  
  2.1. Inequality Regimes and Symbolic Violence ...................... 26  
  2.2. Inclusion .................................................................. 28  
    2.2.1. Definitions of Inclusion .......................................... 28  
    2.2.2. A Revised Definition of Inclusion Based on Legitimacy Theory ........................................... 32  
    2.2.3. Inclusion Training Interventions and Measurements of Inclusion ....................................... 35  
  2.3. Gender-Based Violence ................................................. 37  
    2.3.1. Definitions of Gender-Based Violence ....................... 37  
    2.3.2. Theoretical Grounding of Bystander Interventions for GBV ............................................... 39  
    2.3.3. Sexual Consent and Active Bystander Interventions ......................................................... 41  
  2.4. Moving Forward with Mixed-Methods .............................. 44  
    2.4.1. Social Network Analysis and Inclusion ...................... 50  
    2.4.2. Computational Text Analysis and Trust ..................... 52  
  2.5. Chapter Conclusion .................................................... 54  

Chapter 3: Methodology ....................................................... 55  
  3.1. Field Site Setting ....................................................... 55  
  3.2. Gender-Based Violence Study Research Design ................ 56  
    3.2.1. Participants ........................................................ 57  
    3.2.2. Data Collection .................................................... 59
6.1. Data................................................................................................................................. 141
6.2. Analysis ............................................................................................................................. 146
   6.2.1. Aggregated Scale Analysis ......................................................................................... 146
   6.2.2. Demographic Patterns of Survey #1 ......................................................................... 151
   6.2.3. Demographic Patterns of Survey #2 ......................................................................... 157
6.3. Discussion ......................................................................................................................... 163
6.4. Limitations ......................................................................................................................... 167
6.5. Chapter Conclusion ........................................................................................................... 168

Chapter 7: Quantifying Inclusion with Social Network Analysis ............................................ 169
7.1. Data .................................................................................................................................. 171
7.2. Social Network Analysis .................................................................................................... 173
   7.2.1. Core-Periphery Analysis ............................................................................................ 176
   7.2.2. Centrality Measurements ........................................................................................... 182
   7.2.3. Speaking Times .......................................................................................................... 185
7.3. Discussion ......................................................................................................................... 189
7.4. Limitations ......................................................................................................................... 195
7.5. Chapter Conclusion ........................................................................................................... 196

Chapter 8: Conclusion ............................................................................................................... 197
8.1. Summary of Findings and Contributions .......................................................................... 198
   8.1.1. Research Question #1 ............................................................................................... 199
   8.1.2. Research Question #2 ............................................................................................... 200
   8.1.3. Research Question #3 ............................................................................................... 202
8.2. Implications and Recommendations ............................................................................... 203
8.3. Limitations and Future Research ....................................................................................... 205
8.4. Epilogue ............................................................................................................................. 207

Appendices ................................................................................................................................. 209
Appendix A: Additional Field Site Content .............................................................................. 209
   A.1. Existing Strategies for Mitigating GBV & Fostering Inclusion ..................................... 209
Appendix B: Gender-Based Violence Study .............................................................................. 209
   B.1. Research Design per Cohort .......................................................................................... 209
   B.2. Consent Forms ................................................................................................................ 210
<table>
<thead>
<tr>
<th>Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>B.3. Recruitment Communications and Setbacks</td>
</tr>
<tr>
<td>B.4. Survey Items</td>
</tr>
<tr>
<td>B.5. Student Accommodation Gender Dynamics Focus Group Activity</td>
</tr>
<tr>
<td>B.6. Final Participant Numbers per Survey and per Focus Group</td>
</tr>
<tr>
<td>Appendix C: Inclusion Study</td>
</tr>
<tr>
<td>C.1. Study Communications</td>
</tr>
<tr>
<td>C.2. Consent Forms</td>
</tr>
<tr>
<td>C.3. Surveys</td>
</tr>
<tr>
<td>C.4. Survey Completion Rates</td>
</tr>
<tr>
<td>C.5. Student Responses per Survey and Survey Item</td>
</tr>
<tr>
<td>C.6. Additional Networks</td>
</tr>
</tbody>
</table>

**References** 247
List of Figures & Tables

Chapter 2: Theoretical Overview ................................................................. 25
  Figure 2.1: A New Theoretical Framework of Inclusion/Exclusion and Gender-Based Violence ................................................................. 45
  Figure 2.2: Inclusion/Exclusion Renewed Framework ................................ 47
  Figure 2.3: Gender-Based Violence Renewed Framework ................................ 49

Chapter 3: Methodology .............................................................................. 55
  Table 3.1: Racial and Ethnic Breakdown of GBV Study Participants ........... 59
  Table 3.2: Cohort Participant Numbers Per Research Stage ...................... 62
  Table 3.3: Inclusion Study Research Design for Consent Tiers .................. 67
  Table 3.4: Racial and Ethnic Breakdown of Inclusion Study Participants .... 69

Chapter 4: Quantifying Trust with Mixed Methods ..................................... 81
  Table 4.1: Overview of Vignette Scenarios .............................................. 85
  Table 4.2: Averages & Significance of Bystander Expectation & Permission Scores ................................................................. 94
  Figure 4.3: Likert Scale Responses & Averages per Survey for Bystander Expectation & Permission Items ................................................................. 95
  Table 4.4: Averages and Significance of Community Scale Scores .......... 97
  Figure 4.5: Likert Scale Responses and Averages per Survey for Community Items ................................................................. 98
  Table 4.6: Averages and Significance of Safety & Trust Scores .............. 100
  Figure 4.7: Likert Scale Responses & Averages per Survey for Safety & Trust Items ................................................................. 101
  Table 4.8: Demographics of Survey Respondents ..................................... 104
  Table 4.9: Bystander Expectation & Permission Scale by Demographic Group ................................................................. 104
  Table 4.10: Community Scale by Demographic Group ......................... 104
  Table 4.11: Trust & Safety Scale by Demographic Group ....................... 104

Chapter 5: Quantifying Trust with Computational Text Analysis .................. 113
  Table 5.1: Focus Group and Individual 1:1 Interview Questions to be Analysed ...... 116
  Figure 5.2: Gender Breakdown of Focus Groups ................................... 117
  Figure 5.3: Gender Breakdown of 1:1 Interviews ................................... 118
  Table 5.4: List of Surveys Included in Development of Trust Lexicon ....... 120
  Figure 5.5: Top 18 Stemmed Words from Trust Questionnaires ............... 126
Chapter 6: Quantifying Inclusion with Survey Analysis ................................................. 140

Table 6.1: Overview of MBIE Likert Scale Statements ................................................. 143
Table 6.2: Total Student Survey Completions per School ........................................... 143
Table 6.3: Overall Significance of MBIE Scores and Demographic Groups per Survey .......................................................... 148

Figure 6.4: Likert Scale Responses per Survey for Survey Item #6- Last to Know .......... 150
Figure 6.5: Likert Scale Responses per Survey for Survey Item #4- Invited to Informal Activities ................................................................................................................. 150
Figure 6.6: Likert Scale Responses per Survey for Survey Item #3- Informed About Informal Activities ................................................................. 150

Table 6.7: Significance of Demographics per Survey #1 MBIE Item ....................... 152
Figure 6.8: Likert Scale Responses per Racial Group for Survey Item #1- Share Information ...................................................................................................................... 155

Figure 6.9: Likert Scale Responses per Racial Group for Survey Item #5- Contribute Opinion ................................................................................................................. 156
Figure 6.10: Likert Scale Responses per Racial Group for Survey Item #6- Last to Know ...................................................................................................................... 156

Figure 6.11: Likert Scale Responses per Racial Group for Survey Item #7- Opinion Before Decisions .............................................................................................................. 156

Table 6.12: Significance of Demographics per Survey #2 MBIE Item ..................... 158
Figure 6.13: Likert Scale Responses per Racial Group for Survey Item #1- Share Information ...................................................................................................................... 160

Figure 6.14: Likert Scale Responses per Racial Group for Survey Item #5- Contribute Opinion ................................................................................................................. 161

Figure 6.15: Likert Scale Responses per Racial Group for Survey Item #6- Last to Know ...................................................................................................................... 161

Figure 6.16: Likert Scale Responses per Racial Group for Survey Item #11- Outside Meeting Invites .............................................................................................................. 161
Chapter 7: Quantifying Inclusion with Social Network Analysis

Table 7.1: Total Student Attendance per Tutorial Recording
Figure 7.2: Racial Demographics of Tutorials (Not Including Tutors)
Table 7.3: Example of an Adjacency Matrix (Tutorial C-Recording #2)
Figure 7.4: Example of an Inclusive Network
Figure 7.5: Cumulative Network for Tutorial A
Figure 7.6: Cumulative Network for Tutorial B
Figure 7.7: Cumulative Network for Tutorial C
Figure 7.8: Start and End of Term for Tutorial A with MBIE Scores
Figure 7.9: Start and End of Term for Tutorial B with MBIE Scores
Figure 7.10: Start and End of Term for Tutorial C with MBIE Scores
Table 7.11: Centrality Measurements for Tutorial A
Table 7.12: Centrality Measurements for Tutorial B
Table 7.13: Centrality Measurements for Tutorial C
Table 7.14: Top Five Speakers per Tutorial (Minutes)
Figure 7.15: Total Speaking Times per Racial Demographic Group
Figure 7.16: Breakdown of Speaking Time per Racial Group (No Tutors)
Table 7.17: Top Five Speakers per Tutorial (Minutes) with Interruptions
Figure 7.18: Interruptions Administered per Racial Demographic Group
Figure 7.19: Interruptions Received per Racial Demographic Group

Appendix

Table B.1: Gender-Based Violence Research Design per Cohort
Table B.6a: Cohort Participant Numbers per Survey Group
Table B.6b: Cohort Participant Numbers per Focus Group
Table C.4: Total Student Survey Completion Rates (%) per Survey and Tutorial Group
Figure C.5a: Likert Scale Responses per Survey for Survey Items #1-#4
Figure C.5b: Likert Scale Responses per Survey for Survey Items #5-#8
Figure C.5c: Likert Scale Responses per Survey for Survey Items #9-#11
Figure C.6a: Mid-Term Network for Tutorial A (Guessed Demographics)
Figure C.6b: Mid-Term Network for Tutorial A (Non-Guessed Demographics)
Figure C.6c: Mid-Term Network for Tutorial B
Figure C.6d: Mid-Term Network for Tutorial C ......................................................... 245
Figure C.6e: Cumulative Network for Tutorial A (Non-Guessed Demographics) ...... 246
Chapter 1: Introduction

“[...] if I bring to the class only analytical ways of knowing and someone else brings personal experience, [...] then I humble myself and respectfully learn from those who bring this great gift. [...] fundamentally I believe that combining the analytical and experiential is a richer way of knowing.”
- bell hooks (1994, p. 89)

The initial idea for this piece of research struck me during my master’s programme. I remember the day well. It was a sunny late summer day which meant I was uncharacteristically willing to head the 1.5 hours on an oppressively hot subway from my home to campus for an extracurricular event. Looking back, the fact that the event was catering free Mexican food may have been my main reason for attending. I went to the event alone, loaded up on Dos Toros, and sat down feeling at peace surrounded by other Latinx as well as Black postgraduate students. The panel we were there to listen to was led by a group of four university professors who each identified as Black, Asian, or Minority Ethnic (BAME\(^1\)). It was one professor, Dr. Courtney Cogburn, who posed a question that may have single handedly launched the writing of this thesis. She asked us, a group of almost 40 or so highly educated postgraduates at one of the most elite universities in the U.S., “How is racism in the air you breathe?”. Not one of us had an answer.

This research is concerned with the social phenomena we all know exist, that some of us experienced firsthand, but that is, at times, hard to see. It is subsequently difficult to quantify. Most notably, the last five years have borne witness to a resurfacing gender-based violence (GBV) movement and a growing diversity and inclusion (D&I) movement to counteract some of the damaging effects of these social phenomena. Following the 2020 U.S. deaths of Ahmaud Arbery, George Floyd, and Breonna Taylor, nations worldwide saw a reignited awareness of and support of the Black Lives Matter movement. This opened up larger conversations about D&I across social identities, not just race. Then, in 2021, Britain experienced a national reckoning with gender-based violence upon the murder of Sarah Everard by a police officer in London. In the initial stages of this research, I mainly focused on inclusion/exclusion.

\(^1\) Throughout this dissertation, I am choosing to refer to people who identify as non-white as ‘BAME’. This is despite the UK Government advocating for a discontinuation of the phrase ‘Black, Asian, and Minority Ethnic’, or ‘BAME’ for short. While BAME appropriately avoids centring whiteness, it places emphasis on some ethnic minorities more than others. While it would be most ideal to be precise in the language I use by referring to specific racial and ethnic minority groups, this is not always possible given the statistical weight needed for analysing my study’s sample sizes. Additionally, ‘BAME’ is still the term utilised in official documents at the field site studied. Therefore, when possible, I refer to individual racial and ethnic groups. However, the vast majority of the time, I refer to ‘BAME’ as one conjoined group. It is not my intent to erase individual group’s narratives and experiences in doing so. Rather, I view this as an unfortunate limitation to the analysis of this smaller scope of work.
However, over the course of the past three years, the scope expanded to include gender-based violence too. Exclusion and gender-based violence are generally only seen by others in their most extreme forms: outright discrimination, deaths, hate crimes, and rape. What is more difficult to see is that when they are rendered invisible, normalised in the air we breathe, they still cause harm on a global scale. For this reason, there is an urgent need to measure and to mitigate them.

This thesis advocates for the radicalisation of inclusion/exclusion and gender-based violence research, by borrowing sociological theories that advance the critiques of inequality within organisations like higher education institutions (HEIs) (Romani et al., 2021). According to Bourdieu (1987), social reality is created by invisible relationships between individuals who occupy social spaces dependent on their cultural, symbolic, economic, and social capital. This social stratification, and individuals’ corresponding access to capital, are influenced by one’s intersecting identities (Crenshaw, 1991). Identities in this piece of research are likened to dimensions of diversity; the seen and unseen characteristics of people that differentiate individuals from one and another (Mor Barak et al., 2016). Seen characteristics are usually an individual’s gender, race, ethnicity, and age, among others (Mor Barak et al., 2016). Unseen characteristics could include, for instance, a person’s past educational experiences, nationalities, neurodiversity, etc. (Mor Barak et al., 2016). Thus, what constitutes identity is vast, compounding, and socially constructed. Dimensions of diversity may be acquired from experiences over time, or exist as traits maintained since birth (Gaither, 2017).

Regardless, each dimension of diversity may influence a person’s access to capital, and in turn, influence the social space they occupy within a socially constructed hierarchy.

The setting of this dissertation’s research is an organisation, and more specifically a UK higher education institution (HEI). Organisations are worthwhile focal points to study inclusion/exclusion and gender-based violence because, like any organisation, they maintain structures that create norms and routines (Amis et al., 2020). Sequentially, these structures and their norms and routines work to organise people to get work done. At the same time, they can inadvertently prop up some groups of similarly identifying people above others within the organisation’s socially constructed hierarchy (Amis et al., 2020). This demonstrates how organisations provide valuable starting points for research into inequalities. Furthermore, higher education in the United Kingdom (UK) is diversifying away from the quintessentially young, white, able-bodied male student. According to the UK Higher Education Statistics Agency (HESA) (2022), 57% of all higher education students were female, while 73% of all enrolled students identified as white in the academic year 2021-2022. The percentage of students who identify as white and the percentage of students without disabilities both decreased by 1% from 2021 to 2022 (HESA, 2022). The number of first year students over the age of 30 is also continuing to increase over time (HESA, 2022). It is
vital to understand how these changing demographic groups are being affected by inclusion/exclusion and gender-based violence. To investigate these social phenomena within organisations, I have chosen to research one higher education institution in particular.

Building on the concept of social hierarchies and stratification are inequality regimes. An inequality regime emerges when an organisation adopts, reconstructs, and reinforces pre-existing social hierarchies (Acker, 2006). This results in an organisation maintaining its own invisible hierarchical status quo usually reflective, to an extent, of local, regional, and national inequalities (Acker, 2006). An organisation’s inequality regime consequently pulls from, but does not perfectly reflect, the larger social stratification orders in which it operates (Tomaskovic-Devey & Avent-Holt, 2019). An example of this is a study of two tortilla factories owned by the same parent company with the only differentiator being that one was located north of the Mexico-US border and the other a few kilometres south (Tomaskovic-Devey & Avent-Holt, 2019). In the American factory, institutional inequalities were influenced by immigration status, while in the Mexican factory, inequalities were based on skin-colour and regional origin (Tomaskovic-Devey & Avent-Holt, 2019). With that, certain salient dimensions of diversity from pre-existing social hierarchies may be similarly salient in one institutional inequality regime, but less relevant in another (Armstrong et al., 2011; Crenshaw, 1991). Accordingly, inequality regimes, while contextual, are unequal and oppressive structures that dictate an individual’s (in)access to power within an institutional context (Acker, 2006).

Inequality regimes are highly varied, not just because they are dependent on the larger societal context the organisation is found (such as in the tortilla factory example), but also because organisational interventions targeting organisation-wide behavioural change can alter an internal inequality regime. That being said, individuals do have some agency to modify their institution’s inequality regime (Abdelnour et al., 2017). Groups of individuals can act as change agents within an organisation by altering their behaviours, and subsequent policies, accountability measures, strategies, routines, norms, technologies, etc., that all may influence their organisation’s inequality regime (Abdelnour et al., 2017; Acker, 2006; Tomaskovic-Devey & Avent-Holt, 2019). Yet, as much as an organisational intervention may empower individuals to mitigate hierarchical patterns of oppressive power and domination, there is a chance these wide-scale interventions do very little to the existing inequality regime in the long-term (Dobbin & Kalev, 2018; Evans et al., 2019). From a sociological perspective, this makes researching organisational interventions critical to understanding how social stratification of an inequality regime may shift within an organisation according to different dimensions of diversity.
Researchers, like myself, interested in inclusion/exclusion and gender-based violence within institutions would be remiss to continue ignoring the gap in the literature of studying inequality regimes and organisational interventions. Studies of organisational interventions affecting culture, and thus inequality regimes, may include research into hiring practices, structural changes, compensation, role allocation, and promotions (Amis et al., 2020). Culture in this case may be defined as the amalgam of beliefs, communications, values, and norms that direct a group of people (Ahmed & Shafiq, 2014). Another example of an evidence-based organisational intervention aimed at affecting culture is training for general group members and leaders alike (Lacerenza et al., 2018). My personal motivation to study training interventions’ potential effect on inequality regimes in higher education is both due to the gap in the literature, as well as the mixed results of inquiries into the effectiveness of these interventions (Bonar et al., 2022; Dobbin & Kalev, 2018; Htun et al., 2022; Evans, 2019; Symeonidou, 2017). Despite the mixed results, there is a growing popularity of D&I and gender-based violence trainings within academia (Dobbin & Kalev, 2018; Universities UK, 2019).

Davidson & Ferdman (2001, p. 36) put it best that “to choose not to engage in dialogue about diversity in almost any modern organization is just plain dumb”, and higher education institutions are (usually) not dumb. Conversations in education around inclusion as we define it today began in the 1990s (Symeonidou, 2017). Although the conversation originally was preoccupied with special education systems and students with disabilities, inclusion in HEIs now concerns all students from all backgrounds (Forlin, 2010). To spearhead D&I efforts, many organisations, HEIs included, implement D&I training interventions. American organisations spent $8 billion on such trainings in 2017 alone, while global spending on diversity and inclusion is expected to grow to $24.3 billion by 2030 (Kirkland, 2017; Global Industry Analysts, 2022). According to one Universities UK (2023) survey of their membership of 140 universities in England, Northern Ireland, Scotland, and Wales, all survey respondents had some kind of unconscious bias training in place for both staff and students. Despite the rise in popularity of and financial investment in D&I trainings, there is little agreement within educational institutions of how to achieve inclusive education (Florian, 2019). The previously mentioned elephant in the room is that inclusion interventions, often branded as a fix-all, also have mixed results in achieving inclusion (Dobbin & Kalev, 2018; Symeonidou, 2017). Additionally, there’s a clear discomfort in terms of addressing certain demographics to be included in HEIs, especially when it comes to race in the UK (Bhopal & Henderson, 2021).

Higher education institutions are also increasingly concerned with gender-based violence, and more specifically, sexual violence (Hirsch and Khan, 2020). Despite the popularity and just the same with D&I training interventions, sexual consent and bystander education training programmes present mixed results in combating gender-
based violence (Bonar et al., 2022; Evans et al., 2019; Htun et al., 2022). A National Union of Students survey found that one in four students in the UK experience unwanted sexual behaviour showing the need for university intervention (Phipps & Smith, 2012). Yet, this statistic could be anticipated to an extent by the fact that 48% of students surveyed in another UK study felt unprepared by their secondary schools for sex and relationships at university (Hillman, 2021). The #1 reported impact of a lack of sex education in school was being unsure of what consent is and how to navigate it (MacDougall et al., 2020). For that reason, universities are turning to these sexual consent and bystander interventions as one way to counteract these staggering statistics. According to another Universities UK (2019) survey of their membership, 65% of respondents are actively delivering student consent training and 59% are actively delivering student bystander training. Training programmes are found to be most effective when participants are young, but older than 15 years old (Piolanti & Foran, 2021; Katz & Moore, 2013). In fact, one study found formal sexual consent courses in secondary school significantly reduced university students’ risk of penetrative assault (Santelli et al., 2018). This inevitably puts universities at the forefront of who should be administering sexual consent and bystander education courses, especially keeping in mind those students who lacked consent education or received a fragmented sex education prior to university enrolment.

This research aims to act as a proof of concept by studying two training interventions at one UK higher education institution. In doing so, I will also put forth a new theoretical framework for gender-based violence and inclusion/exclusion. As I will detail further in my theoretical overview chapter, there is a gap in the literature whereby evaluations of training interventions are limited in quantity, in consistency of results, and in diversity of methods. For the evaluations that do exist, researchers to date over-rely on survey methodologies to understand how training interventions affect inclusion/exclusion and gender-based violence. While critical in providing insight into one’s personal experience of these two social phenomena, this methodology overlooks more group-level and harder to see manifestations of inclusion/exclusion and gender-based violence. My new theoretical framework overcomes the limitation of a purely survey methodology by putting emphasis on more reproducible forms of group-level analysis.

In adopting this theoretical framework, my research thus innovates both the inclusion literature and gender-based violence literature by incorporating computational social science methods into a mixed-methods research design. I specifically show how group-level analysis of more covert manifestations of these social phenomena can leverage social network analysis and computational text analysis. Furthermore, by taking a sociological approach to this research, I hope to provide a critique of power, hierarchy, and inequalities that reside within one academic institution’s inequality.
regime. In this way, while I will not conduct a purely pre-test and post-test experimental research design, I do hope to add to the literature by investigating the inequality regime after one inclusion training intervention and one gender-based violence training intervention. I strive to help researchers and practitioners alike better understand how to finetune training programmes aimed at mitigating the more difficult to see manifestations of these social phenomena. As a researcher and as per bell hooks (1994), my ultimate intent in carrying out this research is to be humbled and to learn, all the while combining the experiential with the analytical.

1.1: Research Aims and Questions

While I explore inclusion/exclusion and gender-based violence through two studies, my two aims are unified across both pieces of research. Thus, the overall goals of this research are two-fold. In this thesis and through my empirical analyses, I seek to make a sociological theoretical contribution and methodological contribution to the literature. The first is through my new theoretical framework which focuses on the duality of gender-based violence and inclusion/exclusion as either group- or individual-level and as ranging from covert to overt. The latter methodological contribution is through my use of computational social science methodologies in combination with survey and thematic analyses.

First, I acknowledge that higher education institutions, as any institution, are fundamentally not neutral grounds (Amis et al., 2020). Institutions maintain their own inequality regimes and cannot exist without inequalities (Acker, 2006; Amis et al., 2017). With that, inequalities are normalised (Amis et al., 2020). They are maintained as part of everyday actions, counterintuitive to our ideas of grandiose acts of power, oppression, and domination (Amis et al., 2020). Inclusion/exclusion and gender-based violence, whether covert or not, therefore can and should be studied as part of inequality regimes. This is why my first two research questions strive to answer how training interventions employed by one higher education institution shift the inequality regime in place. In this way, inequality regimes are studied here in terms of inclusion/exclusion and trust of students to be protected from gender-based violence. Power, oppression, and domination within the institutional and larger societal context will be central in my analyses. The first two research questions are empirical in nature and read as follows:

1. What variations in inclusion/exclusion are associated with the implementation of a diversity and inclusion training intervention within a university classroom?
2. What variations in social trust and institutional trust within a university’s student population are associated with the implementation of a gender-based violence training intervention?
As stated, while my empirical findings are important in this dissertation, my overall research aim is theoretical and methodological. Computational social science is increasingly popular in sociology studies, but these methods remain underutilised in the inclusion/exclusion and gender-based violence literature. In this research, I employ both social network analysis and computational text analysis not just to explore how these techniques can be used, but also to revitalise the inclusion/exclusion and gender-based violence research. I attempt to show how my proposed theoretical framework can help move the literature beyond that of the individual experience. Consequently, it allows for more reproducible measures to be incorporated. My third and final research question hopes to combine the experiential and analytical to show the power of mixed-methods computational research by asking:

3. Based on the proposed theoretical framework for gender-based violence and inclusion/exclusion, can a research blueprint be established for more holistically measuring trust and inclusion utilising computational social science methods?

1.2. Positionality and Reflexivity

Critical to my research were the concepts of positionality and reflexivity. Positionality is understanding who I am and how my identities, ways I view the world, and knowledge influence my research process (Corlett & Mavin, 2019). Reflexivity takes positionality a step further whereby I self-monitor and self-reflect throughout the process (Corlett & Mavin, 2019). For me, this meant that although I designed the research process and I can be seen as an owner of this research, I am also an active part of it given the identities I hold. At a high-level, I’m a student, a female living in a male-dominated world, a D&I professional, a graduate of well-funded Global North educational institutions, a gender-based violence victim-survivor, an owner of two hearing aids, a foreigner, and a white-passing ethnic minority navigating a predominantly white institution. All these identities make me emotionally connected to my work and influence the way I understand what inclusion/exclusion and gender-based violence are.

Furthermore, my positionality directly impacts what my intentions are. I cannot claim to be neutral in this research process, especially after over three years of all-encompassing PhD work. At the end of the day, I helped implement the gender-based violence intervention and co-designed the inclusion intervention; I would like to see

---

2 I use the term ‘victim-survivor’ to highlight the duality of ongoing harm and triumph in one’s lifetime after sexual violence. While this is the term I feel most aligned to and utilise throughout this dissertation, I would like to emphasise that some may prefer the term ‘survivor’ and others may prefer the term ‘victim’. 
these interventions do something positive. Related to this, I cannot separate the fact that I have been a D&I professional for the last six years from my research. Prior to academia, I spent two years working in-house in the private sector on diversity and inclusion teams, and one year as a D&I associate consultant working mostly with start-ups and non-profit organisations. I’ve managed 40+ employee resource groups, created and conducted D&I trainings, revamped entire recruitment processes to be more inclusive, and orchestrated interviews with D&I stakeholders from entry-level employees to board members. My practitioner point of view has shaped my research as a scholar and as much as theory is at the heart of this thesis, so too are my thoughts on the practical. In each of the four findings chapters, regardless if I am elaborating on the theoretical or methodological contributions of my work, the practical implications for practitioners trickle in. My positionality inevitably solidifies me to be equally concerned with how my research shapes what practitioners are doing on the ground, not just how it will shape the literature.

Related to my positionality and reflexivity, I would like to take a moment to expand on my ontological and epistemological position. I identify as a pluralist, whereby I maintain an appreciation for both the objective and subjective. This is evident in my theoretical framework proposed in Chapter 2. I claim that research that does not study both perceptions and observable behaviours is important, but intrinsically limited. Note though that of my four research methods, thematic analysis is my only qualitative work. My computational text analysis pulls from qualitative data. Consequently, I found that I struggled epistemologically as other mixed method researchers have (Gobo, 2023; Hall, 2013). To date though, I follow some mixed-methods researchers who utilise computational methods in that I am drawn the most to critical realism. In critical realism, an objective reality exists (O’Mahoney & Vincent, 2014). At the same time, perceptions dictate a person’s reality (O’Mahoney & Vincent, 2014). This makes context, both current and historical, a priority for critical realists (Shannon-Baker, 2016). An individual’s perception in one context may affect their behaviours and perceptions in another context. In my own research, while I am studying the university experience, it cannot be detached from the larger context in which the university is situated. What is happening, and has happened, in UK education in terms of inclusion/exclusion and gender-based violence must be taken into consideration throughout my findings. Readers will undoubtedly see my preoccupation with the larger social context, beyond the one university I study, emerge in my writing.

While I do not have enough word count to list all the ways I have been part of and influenced my research, I do want to share a few ways I have been an “insider-outsider” (Bukamal, 2022, p. 328). Bukamal (2022, p. 328) coined the phrase “insider-outsider ambivalence” to conceptualise how researchers fluctuate between being part of and dissociated from their own work, which in turn impacts their credibility, approachability,
and rapport with participants. I strongly believe in the concept, but disagree with the word ‘ambivalence’. It creates an idea that we as researchers do not care whether we are insiders or outsiders. On the contrary, I care immensely how my experiences as an insider-outsider have impacted my research, my participants, and my findings. To illustrate this, the extent at which student participants interacted with me highly depended on their role in the research. Some students I never met; I live as a faceless name in their email inbox. Others I met in person or through video. They saw my face, heard my voice, experienced my self-deprecation, and saw that I am covered in tattoos. One participant even met my lovely dog, Lottie, when Lottie decided to rummage through the bin midway through our Microsoft Teams interview. It would be impossible to say how these more personal interactions influenced how participants responded to the research; but I assume that meeting me made me more of an insider to them. Related to this, I was more of an insider for tutors than students; I met them all in-person, and just like a lot of them, I am a tutor, a PhD student, and I care deeply about my students’ experiences. How comfortable and ‘insider’ I felt with participants also impacted the way I dressed. I found myself attempting to be a slightly more elevated and relatable graduate student, but not so stuffy that it drew attention to my presence or that it formalised the occasion. This was especially the case in the focus groups and observations. There are infinite other ways I felt my insider-outsider status shift throughout this process, and there is no way I can perfectly describe how my credibility, approachability, and rapport shifted along with it.

1.3: Structure of the Thesis

Chapter 2 situates my work in the academic literature by providing a theoretical overview. It discusses which theorists I build on to conceptualise a theoretical framework for my own research. Due to the interdisciplinary nature of my research and my own educational background, I draw on research from sociology, organisational development, education, business management, social psychology, and computer science. To start, I build on the concept of inequality regimes, first introduced in this current chapter, to prove why gender-based violence and inclusion/exclusion are products of inequality regimes. I then follow with a discussion of the inclusion/exclusion and gender-based violence literatures. I focus on van Dijk & Khattab’s (2021) legitimacy theory of inclusion and Banyard’s (2011) ecological model of bystander intervention to expand the measures possible for inclusion/exclusion and gender-based violence, respectively. I further explore how current measures of inclusion/exclusion and GBV are highly subjective by solely investigating the personal experience. I showcase how incorporating my new theoretical framework for these concepts allows researchers to engage with more reproducible measures through
computational social science techniques. I emphasise that the subjective is necessary to understand one’s personal experience, but that computational social science offers an additional analytical lens through which to view group-level and more difficult to see inclusion/exclusion and gender-based violence phenomena. I end this chapter explaining how computational social science methodologies are already being used to do just this, and how I plan to build on past studies.

Chapter 3 provides an in-depth explanation of my methodology, while my findings run from Chapter 4 through Chapter 7. Given this research is split into two separate studies, Chapter 3 first discusses the gender-based violence research study and then the inclusion/exclusion research study. For each study, I recount my data collection and my subsequent data analysis. I provide an extensive overview of how I combine survey methodologies with computational social science methodologies like social network analysis and computational text analysis. Chapter 4 and Chapter 5 both address social trust (student-to-student trust) and institutional trust (student-to-institution trust), but with different methodologies. Chapter 5 and Chapter 7 are both computational and can be seen to be paired with the non-computational Chapter 4 and Chapter 6, respectively. I intentionally split my methods between computational social science methodologies and non-computational methodologies to show how different research designs can build on each other’s conclusions.

As a result, Chapter 5’s findings build on Chapter 4’s findings. Chapter 4 utilises thematic analysis of focus group data and some survey analysis to understand how self-trust, social trust, and institutional trust change before and after the consent and active bystander training. Five themes emerge, but the main finding is that social trust is higher than institutional trust. Within social trust, nuances relating to gender dynamics, environmental factors, social relationship dynamics, and overall gender-based violence knowledge and awareness also emerge. I put forth practical suggestions to address these nuances, while I also highlight that students post-training place a significantly higher importance on safety with peers than students pre-training. Chapter 5 uses a computational text analysis approach to address the same question of trust as it relates to gender-based violence. In this chapter, I develop a bespoke trust dictionary through word embeddings to quantify levels of social trust and institutional trust in the same focus group data and additional 1:1 interview data. Chapter 5 is in agreement with Chapter 4 by reaching the same conclusion as Chapter 4 in that social trust is higher than institutional trust. Therefore, Chapter 4 can be seen as a standalone empirical contribution. Chapter 5 provides a methodological and theoretical contribution offering exactly how a computational method can provide a more reproducible way to verify findings from a qualitative method. In both chapters, I extensively discuss how inequality regimes persist after the consent and active bystander training and I ground the conclusions in practical next steps for the university. The combination of Chapter 4
and Chapter 5 share how the new theoretical framework can be applied to investigate group-level covert manifestations of gender-based violence and operationalised through mixed-methods research, including computational text analysis.

Chapter 6 and Chapter 7 both apply the new theoretical framework to investigate inclusion at the group-level. Chapter 6 is the first study to adapt the Mor Barak Inclusion-Exclusion scale to a classroom setting. Using statistical analysis, I find a lack of concrete evidence for an overarching inequality regime by any one demographic group in the classrooms studied. There is evidence though that race and ethnicity are the most salient dimension of diversity when it comes to differences in perceptions of inclusion. Furthermore, inclusion scores were significantly higher later in the term on Survey #2 than earlier in the term on Survey #1. At the same time though, Chapter 7 offers another lens through which to view more covert inclusion/exclusion within the classroom by utilising social network analysis. In doing so, it offers a counter to Chapter 6’s conclusion by finding evidence of an inequality regime when it comes to race and ethnicity, especially in terms of participation and contribution. Chapter 7 directly links this finding to the debate around international students in UK higher education institutions. Together, these chapters offer a way to monitor inclusion within the classroom and create a feedback loop for universities hoping to better engage all students. With that, both chapters showcase how the university can further refine future inclusion training interventions.

Chapter 8 concludes this thesis by reiterating the empirical, theoretical, and methodological contributions of this work while offering a way forward. Empirically, I write how Chapter 4 and Chapter 5 together quantify the different levels of trust after sexual consent and active bystander training. Chapter 4 further highlights how an inequality regime may exist in terms of trust for students identifying as a sexual and gender minorities. As a result, in terms of Research Question #2, thematic analysis points to social trust being higher than institutional trust. Chapter 6 and Chapter 7 show how a race and ethnicity inequality regime may be altogether missed in inclusion research if solely surveys are used as the research method. Research Question #1 is therefore addressed by showing that while perceived inclusion is significantly higher in Survey #2 than in Survey #1, this is not true for everyone nor does it translate to observed behavioural inclusion. Race and ethnicity prove to be the most salient dimensions of diversity for differences in behavioural and perceptions of inclusion. Theoretically and methodologically, my four findings chapters show how my new theoretical framework helps move the inclusion and gender-based violence literature forward. In expanding into the group-level and covert manifestations of inclusion and gender-based violence, computational social science can be combined with more subjective forms of analysis to offer a more holistic and reproducible study of these
social phenomena. For Research Question #3, I confirm a blueprint for incorporating computational social science into my theoretical framework.

After reviewing these contributions, Chapter 8 reiterates that organisational interventions should be highly tailored to the inequality regime which they are striving to tackle. Furthermore, by taking the findings from Chapter 4 through Chapter 7 and adjusting the training interventions accordingly, only then can the institution more effectively attempt to mitigate exclusion and gender-based violence. At the same time, my final thoughts on this dissertation’s matter are cynical at best. I stress that the trainings must acknowledge the larger social orders and contexts in which they operate, especially as UK higher education institutions shift to "Higher Education, Inc." (Seal, 2018). I conclude with the question of whether institutions, and especially higher education institutions, can ever be separated from inequalities when they were founded on the principles of and continue to operate within a white-supremacist, neoliberal capitalist, and patriarchal society.

Chapter 2: Theoretical Overview

In this chapter, I conduct a literature review with a foundation in sociology. I also add in theories and methodologies from organisational development, education, business management, social psychology, and computer science. In the first section, I establish how Acker’s (2006) concept of inequality regimes is controlled by three forces derived from Tomaskovic-Devey & Avent-Holt’s (2019) relational inequality theory. Furthermore, I share how gender-based violence and inclusion/exclusion are products of inequality regimes, but rely on different dimensions of diversity as the source of inequality. I emphasise that inequality regimes are highly variable and contextual, meaning that institutional interventions addressing them must be too.

In the next two sections, I dive into the inclusion and the gender-based violence literatures to show how current research is limited by an overemphasis on individual perception. Given the majority of inclusion research has been done in the organisational development and business management literature, I provide an overview of two leading inclusion frameworks by Shore et al. (2011) and Roberson (2006). I critique the frameworks by anchoring my work in van Dijk & Khattab’s (2021) more nuanced legitimacy theory of inclusion. This theory of inclusion maintains that there is both subjective and objective inclusion. I discuss how current measurements of inclusion fail to acknowledge objective, observable behaviours of exclusion and instead over-rely on individual’s own perceptions. I write about how current inclusion training interventions have mixed results in the educational inclusion space (Dobbin & Kalev, 2018; Symeonidou, 2017). I call for an adoption of the legitimacy theory of inclusion to be applied to educational inclusion research in the hopes that post-training measurements
of inclusion interventions will more accurately reveal covert inequality regimes. Following this, I discuss what gender-based violence is and how just the same as inclusion, it is a product of power and hierarchy. I similarly write about how gender-based violence interventions in education are largely effective for short periods of time and only minimally (Bonar et al., 2022; Htun et al., 2022; Evans, 2019). I build on Banyard’s (2011) ecological model of consent and bystander training to revitalise gender-based violence research and training measurements. In doing this, I expand measures post-training intervention to include community-level factors such as social trust and institutional trust.

After reviewing the literature, I showcase my own theoretical framework for researching inclusion/exclusion and gender-based violence. Reminiscent of inequality regimes, this framework discusses how these social phenomena vary in terms of how obvious they are to those who are not experiencing them directly. I explain how these social phenomena also occur both at the individual and group-level. I provide ample examples of my theoretical framework to emphasise how the group-level is understudied due to the limitations of current theories of inclusion/exclusion and gender-based violence. Moving beyond the current overemphasis on the more overt and individual manifestations, I share how I will focus on the group-level and harder to see manifestations of these social phenomena to add to the literature.

From there, I introduce computational social science techniques including social network analysis and computational text analysis. I urge social scientists to consider how these techniques can provide innovative lenses through which to study training interventions. For example, social network analysis can be used to investigate group-level and more covert acts of inclusion/exclusion. Similarly, by acknowledging the group-level and more covert forms of gender-based violence, I can utilise computational text analysis to measure ongoing levels of trust within educational communities. Using these computational methodologies in a mixed-methods approach helps both researchers and practitioners better understand how gender-based violence and inclusion/exclusion within inequality regimes are being (re)created in universities. Consequently, my theoretical framework not only moves the literature forward, but also helps provide feedback to calibrate inclusion and gender-based violence interventions in practice.

2.1: Inequality Regimes and Symbolic Violence

Tomaskovic-Devey and Avent-Holt’s (2019) relational inequality theory offers insight into the forces that control inequality regimes. The foundation of this theory is derived from Charles Tilly’s categorical inequality theory. At the crux both theories is that inequality regimes are a result of present-day interactions between groups with
unequal power (Tilly, 2000; Tomaskovic-Devey & Avent-Holt, 2019). In relational inequality theory, the three defining and active forces that influence modern day social stratification are exploitation, social closure, and claims-making (Tomaskovic-Devey & Avent-Holt, 2019). These forces occur in several power relation arenas, including families, work, politics, and of course, most relevant to this research, education. Regardless of context and just as with inequality regime’s structures, the forces of relational inequality theory can be very subtle. The theory thus shows how overt discrimination persists (ex. Hate crimes), so too does covert discrimination (Tomaskovic-Devey & Avent-Holt, 2019). Ultimately, it is the covertness of these forces that render inequality regimes almost invisible. The mechanisms can also be enacted subconsciously or without malintent (Tilly, 2000). This subtlety makes it difficult to counteract, monitor, and measure inequalities, and their resulting violence, at scale. It is clear though that relational inequality theory offers a powerful lens for critiquing inclusion/exclusion and gender-based violence within higher education discourse; forcing universities to investigate the different forces that keep their unique inequality regimes in place.

Gender-based violence can be analysed in terms of claims-making and exploitation. Claims-making is the political and relational process whereby actors in a given social order will attempt to declare resources for themselves, but only some will do so successfully (Tomaskovic-Devey & Avent-Holt, 2019). The effectiveness of an individual’s claims-making is dependent on their categorical group, their relative power, and their persuasiveness (Tomaskovic-Devey & Avent-Holt, 2019). A gender-based violence example of this is that while sex may seem like a private act, it operates within a larger public societal sphere (Olufemi, 2020). When it comes to binary and non-binary genders, there are clear power hierarchies and gender roles at play (Burkett & Hamilton, 2012; Olufemi, 2020). Power relations do not dissipate in sex when critiquing who can make claims on or exploit whose bodies (Olufemi, 2020).

Inclusion most directly relates to social closure. Social closure consists of both opportunity hoarding by in-groups and exclusion of out-groups by in-groups (Tomaskovic-Devey & Avent-Holt, 2019). Outright exclusion is becoming increasingly unacceptable due to social stigma and legislation, but opportunity hoarding is still, at times, culturally acceptable (Tomaskovic-Devey & Avent-Holt, 2019). A higher education example of this may be that universities cannot outright discriminate against historically marginalised students anymore. The UK Equality Act 2010 now protects all students, and prospective students, from discrimination regardless of identity (Koutsouris et al., 2022). Other forms of exclusion or opportunity hoarding still persist though. As of 2021, one in every three first year international students in the UK is Chinese (Higher Education Policy Institute, 2021). Even with the legal framework protecting from outright discrimination, there are exclusionary and opportunity hoarding
forces that these students must overcome to enrol and during their time as a student. This can involve growing international fees and students having to have the resources to pay them (Bamberger et al., 2019). It can also involve Chinese students having to navigate the independent learning discourse style of UK classrooms, which some students find excludes those who do not participate by speaking up quickly, interrupting, or asking questions (Liu et al., 2022; Wu, 2015). With that, this research adopts a sociological perspective of inclusion, whereby societies maintain “architectures of inclusion [and] landscapes of exclusion” resulting in the maintenance of a hierarchical inequality regime (Allman, 2013, p. 2).

Bourdieu’s (2001) concept of symbolic violence is useful in further understanding how inequality regimes may covertly operate within higher education institutions. Symbolic violence is a system of domination that persists through “imperceptible and invisible” forces (Bourdieu, 2001, 2). Symbolic violence therefore highlights how the three forces upholding inequality regimes are not usually grandiose, grossly intentional, and painfully obvious acts of domination. Rather, symbolic violence operates in run of the mill interactions in a person’s day to day; it is stealthy, subtle, and unconscious (Reay, 2022; Gordan & Zainuddin, 2020). The concept of symbolic violence also sheds light on how inequality regimes exist at the structural- or institutional- level, but how they are maintained also at the individual-level (Reay, 2022). The individual reinforces the structure of an inequality regime through their habitus (Bourdieu, 2001). Conversely, the hierarchical structure in place reinforces processes of domination within someone’s habitus too (Bourdieu, 2001).

This calls into question the role of agency and policy in mitigating inequality regimes. Although educational systems and institutions are sites of inequality reproduction, research finds these often go untouched by policy (Reay, 2022). Yet, Bourdieu asserts the necessity of “collective organisation and action and effective weapons, especially symbolic ones, capable of shaking the political and legal institutions which play a part in perpetuating […] subordination” (Bourdieu, 2001, ix). In consideration of that, this chapter will detail how training interventions targeting inclusion/exclusion and gender-based violence play a central role in attempts to activate one’s agency to dismantle the covert forces that uphold inequality regimes within higher education institutions.

2.2: Inclusion

2.2.1: Definitions of Inclusion

Diversity and inclusion are often used interchangeably as the “language du jour”, but inclusion is in fact its own distinct concept (Gebru, 2020, p. 9; Mor Barak, 2015).
Furthermore, researchers from all relevant fields disagree on an inclusion definition and lack scientific evidence necessary for consensus (Armstrong et al., 2011; Chung et al., 2020; Ferdman, 2013; Kuknor & Bhattacharya, 2021; Roberson, 2006; Qvortrup & Qvortrup, 2018; Shore et al., 2011). This struggle of the inclusion field to move together united has led to the call for “validated, conceptually grounded measures […] Without such scholarship, the inclusion literature will continue to grow and scholars will struggle with needed conceptual clarity” (Shore et al., 2018, p. 186). These attitudes extend to educational research where inclusion research is inconclusive and patchy (Florian, 2019; Qvortrup & Qvortrup, 2018). In one recent study of diversity and inclusion policy documents from the 24 Russell Group universities in the UK, although all 24 reference inclusion, barely any of them actually define what it means (Koutsouris et al., 2022). Some universities equate it to other relevant terms like respect, and another even somewhat laughably defines inclusion as “feeling included” (Koutsouris et al., 2022, p. 884).

Of the numerous and, at times, ambiguous definitions of inclusion, two main definitions have prevailed in the inclusion research literature (Chung et al., 2020; Mor Barak & Cherin, 1998; Roberson, 2006; Shore et al., 2011). The first definition that became prevalent in the academic literature was based on the inclusion-exclusion continuum model by Mor Barak & Cherin (1998), which reflected employee perceptions of their involvement in work groups, influence in decision-making, and access to communication and resources. However, Mor Barak then shifted to Shore et al.’s (2011) model heavily guided by Brewer’s optimal distinctiveness theory, which similarly focuses on an individual’s perception of inclusion. According to Brewer (1991), an individual has opposing needs to feel like they belong in a group and to feel distinct enough from a group at the same time. Too much regard for one’s uniqueness can lead to a sense of isolation and be uncomfortable for the individual desiring to be seen as part of a group (Brewer, 1991). Too much group identity rather than respect for one’s uniqueness is also detrimental because a person can lose their sense of self (Brewer, 1991). Shore et al. (2011) states that only once these two requirements are perceived to be perfectly balanced will an individual feel included. Accordingly, one needs to be respected for their individuality while also being wholeheartedly welcomed into the group. Inclusion is therefore, “the degree to which an employee perceives that he or she is an esteemed member of the work group through experiencing treatment that satisfies his or her needs for belongingness and uniqueness” (Shore et al., 2011, p. 1265).

The other leading definition of inclusion in the academic literature is Roberson’s (2006, p. 21) definition which defines inclusion as “the removal of obstacles to the full participation and contribution of employees in organisations”. This construct instead focuses on collaborative work arrangements and conflict resolution procedures to
increase the number of diverse individuals directly involved in decision-making (Roberson, 2006). Roberson’s (2006) definition of inclusion developed from empirical research on Chief Diversity & Inclusion Officers’ own definitions of inclusion, in an attempt to bridge the practitioner-scholar gap in inclusion research. Other scholars such as Wasserman, Gallegos, and Ferdman (2008) similarly anchor their definitions of inclusion in that of participation and contribution too, rather than solely the experience of belongingness and uniqueness like Shore et al. (2011).

Ferdman (2013) writes that an agreed-upon definition by inclusion researchers at this time is perhaps superfluous, but I argue the lack of an agreed upon definition has inhibited inclusion research. Without an agreed upon definition, researchers are not united in how they are conceptualising and subsequently measuring inclusion. I think both leading definitions could also be extended beyond the boundaries of the workplace as both offer food for thought to inclusion research in the higher education space. More specifically, Roberson’s and Shore et al.’s definitions both can be applied to the higher education context as their focus is on work groups. Shore et al.’s (2011) focus on perceptions of uniqueness and belonging can be applied to a students’ university experience, whether researching a student’s inclusion in a society, a project group, or in the classroom. Inclusion in terms of participation and contribution is reminiscent of the educational concepts of mainstreaming or integration (Hyde, 2017). Tienda (2013) even concludes that inclusion and integration are interchangeable in their meaning. Participation and contribution at a more granular level can also be considered in terms of a students’ participation in campus events or in classroom discussions. Both definitions of inclusion provide a lens through which inclusion research in the higher education space can be explored.

Additionally, I assert any agreed upon definition of inclusion must marry together the idea of one’s individual perception and the behaviours occurring. Below, I outline how Shore et al.’s (2011) definition, with its focus on one’s perception of uniqueness and belongingness limits itself to understanding inclusion through one’s personal experience of inclusion/exclusion. This completely disregards any exclusionary behaviours that occur unless a person perceives them as exclusionary. Roberson (2006) takes a more behavioural approach by focusing on participation and contribution. On the contrary, once operationalised, the definition leads to researchers still solely relying on people’s perceptions of their participation and contribution. Both behaviours and perceptions of inclusion/exclusion must be incorporated into a definition to move the inclusion literature forward.

In this way, the Shore et al.’s (2011) inclusion definition borrows heavily from social psychology and leads the literature to consider an individual’s experience of inclusion (Shore et al., 2018). These social psychological theories contribute to the development of group membership which in turn, directly dictates who is to be trusted.
and who is included in one’s own social network (Ridgeway, 2011). One such social psychology theory researchers rely on is social identity and categorisation theory, which states that individuals will unconsciously sort themselves into groups based on common dimensions of diversity, such as gender or race (Tajfel & Turner, 1986). Of course, intersectionality and how salient identities compound make these emerging in- and out-group dynamics less straightforward. Regardless, social identity and categorisation theory stems from humans’ need for self-evaluation. Individuals validate their own identities, by validating others’ identities, and decide how they perceive others both within and external to their groups (Festinger, 1954; Tajfel & Turner, 1986). Sociologists refer to this process of establishing categories both within human cognition and in social structures as boundary work and framing, respectively (Massey, 2007). Individuals from one’s own group will be framed as warmer and more competent (Massey, 2007; Ridgeway, 2011). Lastly, social identity complexity theory also shows how in-group favouritism is affected by perceived threat, as more diverse groups with compounding identities will be more accepting due to their lower levels of perceived threat (Miller et al., 2009; Roccas and Brewer, 2002).

All together, these social psychological theories help researchers understand why perceptions of inclusion/exclusion may shift according to the emergence of in- and out-group dynamics, and in turn, how inequality regimes persist. Individuals rely on framing to assign materials, power, and resources in inequality regimes that are unreflective of potentially newer, more egalitarian norms (Ridgeway, 2011). For example, race-ethnicity framing can affect how an individual perceives their work group productivity and commitment, as well as opportunities for advancement (Riordan & Shore, 1997). Gender framing is also critical, especially in ambiguous social situations or where individuals are making quick unconscious decisions (Ridgeway, 2011). Although female stereotypes are improving, there is a cultural lag whereby gender framing women negatively in comparison to men in work contexts persists (Ridgeway, 2011). Khattab et al. (2020) expand on these concepts by theorising that minority employees may even self-frame negatively in a way that leads to lower work network utilisation and subsequently, less professional advancements.

As discussed, framing is not as simplistic either when one considers other salient compounding identities beyond gender and race, such as educational background, marital status, age, etc. Additionally, taking into account framing and saliency of different identities in group contexts, it is important to get one’s opinion of how they feel they are being treated dependent on their identity. At the same time, there is more to inclusion than just how someone is perceiving it. Optimal distinctiveness theory, a linchpin of the Shore et al. (2011) theory of inclusion, shows how the most inclusive groups unite individuals giving them a sense of similarity and belongingness while sufficiently recognising their unique characteristics (Brewer, 1991; Roccas &
Brewer, 2002). While it is clear that the above theories behind Shore et al.’s (2011) framework of inclusion are worthwhile to study, it has led the literature to over-rely on an individual's experience of inclusion (Qvortrup & Qvortrup, 2018). For instance, optimal distinctiveness theory and its focus on ensuring an individual feels both unique and a sense of belonging drives researchers to only investigate inclusion in terms of the individual’s perception of uniqueness and belonging.

On the other end of this, the other leading framework of inclusion, Roberson’s (2006) definition, similarly does not fully capture the nuanced nature of inclusion. Roberson’s (2006) definition, which focuses on an individual’s perception of participation and contribution, has led the literature to focus a bit more on the behaviours occurring while acknowledging the group dynamic and social psychology dynamics at play. Yet, just as with Shore et al. (2011), the resulting literature still overlooks what behaviours have occurred and instead focuses on how individuals feel. Methodologically, this translates into resulting measurements that remain solely subjective, relying on an individual’s memory of how much they were able to participate and contribute. This can be inaccurate if when asked about their participation and contribution someone forgets, overestimates, underestimates, blatantly lies, etc. Without monitoring the behaviours that occurred, this once again places an overemphasis on an individual’s subjective report of their experience even though Roberson’s (2006) definition does not call for this.

2.2.2: A Revised Definition of Inclusion Based on Legitimacy Theory

Sparkman (2019) reflects on the fact that without an agreed upon definition, theoretical frameworks for inclusion have endlessly shifted; however, legitimacy theory presents an innovative and nuanced way forward for inclusion researchers by highlighting how exclusion is omnipresent and not always negative. Although most of the academic literature frames cohesive diverse groups as leading to inclusion as a positive, and incohesive diverse groups as leading to exclusion as inherently negative, this thesis instead shares in van Dijk & Khattab’s (2021) legitimacy theory of inclusion. Legitimacy theory highlights a paradox. Paradoxes have been used in the inclusion literature prior to van Dijk & Khattab (2021), but instead focused on the tensions between self-expression and group conformity, safety and comfort, and boundaries and norms (Ferdman, 2017). Van Dijk & Khattab (2021) propose a paradox whereby inclusion is not inherently positive and exclusion is not inherently negative as much of the literature alludes (van Dijk & Khattab, 2021; Periac et al., 2018). Rather than this binary perspective that inclusion is positive and exclusion is negative, legitimacy theory of inclusion proposes that inclusion and exclusion are not contradictory (Armstrong et al., 2011; van Dijk & Khattab, 2021). The two constructs coexist as interrelated
processes and shift in their respective degrees of positivity and negativity (Hyde, 2017; Qvortrup & Qvortrup, 2018; van Dijk & Khattab, 2021).

This paper agrees that to determine the legitimacy of inclusion/exclusion, researchers must examine not just the contextuality and salience of diversity in a given context, but also both behaviours and perceptions. In this way, legitimacy theory, unlike Shore et al. (2011) and Roberson (2006), simultaneously examines both a person’s experience and the objective behaviour occurring, to label inclusion and exclusion as either legitimate or illegitimate in a given context (van Dijk & Khattab, 2021). For example, a student could not be invited to a presentation group meeting or not added to a WhatsApp study group chat. Those are objectively exclusionary behaviours reminiscent of social closure and opportunity hoarding. What is critical to assess therefore is both the objective behaviour and the experience of the person not included in tandem. Do they feel the behaviour was discriminatory due to their skin colour, or, because of their lack of involvement in the class? Only one is likely to be considered an illegitimate exclusionary behaviour, but both the behaviour and perception of the behaviour must be investigated to determine the legitimacy of the behaviour as exclusionary.

Similarly, imagine COVID-19 is rampant again (knock on wood) and a university has issued a mask mandate for all students and staff. It is halfway through the term and a student has never contributed to a tutorial’s class discussion. On one hand, that can be seen as a self-imposed objectively exclusionary behaviour. Yet, a student in that scenario can still feel included and may just identify as an introvert. On the other hand, another student who has a hearing impairment could feel excluded as the mask mandate makes it virtually impossible for them to read lips. In the latter case, the student could be at a loss and incapable of fully understanding class discussions. The university’s mask mandate in this case would be inadvertently reserving the opportunity to learn for students without hearing impairments. Without examining the contextuality and salience of diversity, as well as the behaviours and perceptions of inclusion/exclusion at play, nuance can be altogether missed. This theory holistically assesses the interplay between inclusion and exclusion in upholding an inequality regime and for this reason, legitimacy theory of inclusion sits at the crux of this research.

Incorporating legitimacy theory means acknowledging both the inclusionary and exclusionary behaviours and perceptions of those behaviours that exist to more effectively evaluate whether inclusion/exclusion, if observed and/or experienced, is legitimately positive or negative. It calls for placing an importance on simultaneously observing the behaviours that occur, while asking people how they felt they were included or excluded after the behaviours happened. In this way, legitimacy theory shares how the same behaviour can be interpreted as exclusion or inclusion depending
on the person. While inclusion/exclusion sometimes may be clear cut, it is usually not going to be the case that everyone agrees inclusion was happening, nor that only exclusion was happening. So, while legitimacy theory can be seen as complicating the idea of ever achieving inclusion in totality (if that is even possible), it offers a critical exploration of the nuances of inclusion and exclusion. The individual experience is inherently going to need to be considered to help determine if the inclusionary/exclusionary behaviours are legitimate or illegitimate. After all, both interpersonal and intergroup processes help uphold power dynamics, so the behaviours within these processes should also be studied (Ferdman & Sagiv, 2012). Thus, legitimacy theory of inclusion shows how the literature’s sole focus on individuals’ perceptions of inclusion is insufficient. Researchers should also explore the behaviours occurring, and not disregard the fact that exclusion goes hand in hand with inclusion. Bringing this all together, there is an opportunity to combine Roberson (2006) and Shore et al. (2011) while tying in van Dijk & Khattab’s (2021) legitimacy theory of inclusion. If inclusion/exclusion coexist, as described in legitimacy theory, and inclusion/exclusion both exist as perceptions and behaviours, as Shore et al. (2011) and Roberson (2006) separately point to too, then it is necessary to look at both perceptions and behaviours for both inclusion and exclusion. In addition to this, while it is rare for researchers to anchor inclusion in the middle of the uniqueness/belongingness and participation/contribution debate, this paper argues for a combination of all four tenets. This means inclusion/exclusion can be researched by looking at the participation and contribution behaviours occurring, along with individuals’ experiences of their uniqueness, belongingness, contribution, and participation. This combination of inquiry into behaviour and perception is necessary for researchers to move the inclusion, and exclusion, field and its resulting methodologies forward.

There is only one definition of inclusion in the literature that seemingly acknowledges the two-fold nature of inclusion as both perception and behaviour. Puritty et al. (2017, p. 1101) writes inclusion is, simply put, “how [people] are treated and how they feel”. I do not necessarily want to add another definition to the inclusion literature, but I do feel like Puritty et al.’s (2017) definition can be amended. Inclusion/exclusion can consequently be “how people are treated and behave in terms of contribution and participation, and how they feel in terms of the former, as well as their needs for uniqueness and belonging.” Methodologically, combining Shore et al. (2011) and Roberson (2006) means bringing in both observations of contribution and participation, as well as surveying people’s perceptions of their contribution, their participation, and their needs for uniqueness and belonging. This behaviour and perception combination is crucial for researchers to revamp the inclusion field and its resulting methodologies forward, all the while being able to determine the legitimacy of inclusion/exclusion.
2.2.3: Inclusion Training Interventions and Measurements of Inclusion

Prior to discussing the effects of inclusion training interventions and how these are measured, it is first imperative to define what constitutes an inclusion training intervention. While the increase in D&I in the private sector is clear, what an inclusion training intervention is remains less clear. In my nearly six years as a diversity and inclusion professional, I’ve found that inclusion training interventions can be wide ranging to say the least. To me, they can essentially include any type of training that engages with diversity and inclusion concepts whether that is inclusion, diversity, equity, equality, racism, social justice, privilege, liberation, decolonisation, respect, disability, etc. Usually, I find the most common inclusion trainings to be unconscious bias training, sometimes referred to as anti-bias training (Dobbin & Kalev, 2018). Studies show though that unconscious bias training and explicit bias training alike have weak effects (Dobbin & Kalev, 2018). I personally have also facilitated inclusive recruiting training, inclusive leadership training, and inclusive teaching training. In addition to this, I’ve personally attended antiracism and community organising training, LGBTQIA+ inclusion training, and inclusive language training. I would umbrella all of these types of trainings under the catchall term ‘inclusion training intervention’.

Academic research into inclusion training interventions has lacked (Ferdman, 2013; Ferdman & Sagiv, 2012). Studies that have investigated the effectiveness of inclusion training interventions show heavily mixed results, both in the organisational development and education literature (Dobbin & Kalev, 2018; Symeonidou, 2017). Symeonidou (2017) in a meta-analysis found inclusion interventions have mixed effects on attitudes, knowledge, and skills of student-teachers. It is also uncommon for teachers to even receive inclusion training because of a general lack of training for teachers in general (Armstrong et al., 2011). Those who do receive inclusion training receive it either as an add-on course, or through content infused into other coursework (Florian & Camedda, 2020; Symeonidou, 2017). The mixed results of these programmes do not help alleviate the common view of inclusion interventions as “fluffy”, or surface-level with debatable return on investment (Pillay & Redfern, 2017).

Although only a speculation, the mixed results are perhaps a result of the covertness and variability of inequality regimes. Afterall, if organisations are implementing cookiecutter inclusion training interventions, such as off-the-shelf unconscious bias training, it makes sense that the results are not targeting the specific inequality regime in place (Githens, 2011). Allan (2006, p. 126) expands on this and reflects, “The standards and accountability culture creates closures, but also catches everyone - policymakers, teacher educators, researchers, teachers, parents and children and young people - in a performance, forced to enact a version of inclusion which is merely about tolerance and management of difference and which leads to an constant
reiteration of exclusion”. This poses the question of despite the popularity of these types of trainings, are they actually affecting the covert oppressive structures in place in higher education?

Before explaining how measurements of inclusion should be altered to incorporate the revised definition of inclusion discussed above, it is first necessary to understand current measurements of inclusion. Measuring the impact of D&I interventions on diversity involves monitoring demographics. Depending on the context, this could be any of the seen and unseen characteristics that make up a person’s identity such as race, ethnicity, gender, sexuality, age, educational background, disability, religion, veteran status, linguistic ability, etc. Quantifying diversity consequently involves a quick adding up of how people identify. Measuring inclusion is less straightforward (Sherbin & Rashid, 2017; Winters, 2013). Overwhelmingly, measurements of inclusion rely solely on individuals’ perceptions of inclusion captured by surveys across varying and shifting dimensions of inclusion. For instance, from the organisational development literature, measurements include the Inclusion Index™ measurement (Kurt & Blass, 2010), the Climate for Inclusion Survey (Nishii, 2013), and the Mor Barak Inclusion-Exclusion Scale (Mor Barak & Cherin, 1998). The most utilised measurement is the Mor Barak Inclusion-Exclusion Scale, which despite Mor Barak et al. (2022) acknowledging behavioural and perceptions of inclusion, remains anchored in just contribution and participation (O’Keefe et al., 2020). Each of these methodologies, despite their different definitions and dimensions, remain anchored in individuals’ perceptions. From the education literature, measurements almost exclusively focus on teachers’ experiences of inclusion and include Attitudes Towards Inclusive Education Scale; Sentiments, Attitudes, and Concerns about Inclusive Education Scale; Teacher Efficacy for Inclusive Education Scale (Symeonidou, 2017). None of these measurements currently incorporate both perceptions of inclusion/exclusion and behavioural observation.

A behavioural measurement of inclusion/exclusion grounded in legitimacy theory could be modelled after Goldberg et al.’s (2016) study on embeddedness. The researchers theorise that embeddedness manifests culturally and structurally (Goldberg et al., 2016). Structurally, people can be embedded if they are rooted in well-connected networks (Goldberg et al., 2016). Culturally, people can be embedded if they fit in with their colleagues on a personal level (Goldberg et al., 2016). The researchers use social network analysis of email metadata to understand to what extent employees are structurally embedded, while they use computational text analysis of the same emails to measure cultural embeddedness (Goldberg et al., 2016). The behavioural measurements taken are therefore less susceptible than surveys to bias. A survey approach could involve asking employees how structurally well-connected and culturally connected they feel. Yet, it would be prone to personal bias, halo effect, and
results could also be skewed if employees feel pressured to complete them in a certain way (Kurt & Blass, 2010; Turnbull et al., 2009). By utilising computational social science to study embeddedness, Goldberg et al. (2016) can highlight otherwise overlooked behavioural manifestations of cultural and structural embeddedness.

Goldberg et al. (2016) provide a sound framework for incorporating behavioural measures into research design, but perceptual measures are still critical in capturing the nuances in patterns of inclusion/exclusion. If these behavioural and perceptual measurements are combined and applied to inclusion research, patterns of inclusion could be studied in relational processes like Goldberg et al. (2016). Behaviours of inclusion and exclusion could be pulled from communication and social systems or relational data, while the legitimacy of inclusion could be captured via a survey. Data from communication and social systems, like WhatsApp group chats, course online discussion forums, and email chains, can be collected. Such digital systems offer what may be referred to as digital traces or artefacts, byproducts of everyday actions and behaviours such as who is and is not invited to meetings, who is communicating with who, or who is and is not copied on emails (Salganik, 2018). Relational data could come from classroom or group project meeting interactions. This would eliminate the overreliance on surveys to measure inclusion, which some researchers have already called for (Ostroff & Bowen, 2016). It would also more purposefully investigate how inequalities occur within day-to-day interactions in an inequality regime, just as Goldberg et al. (2016) did with their study on embeddedness (Acker, 2006; Ferdman, 2003).

Mercer (2010) found that dialogic patterns in the classroom show how power is negotiated, how information is exchanged, and how classroom dynamics are constructed. When paired with a subjective measurement of inclusion such as a survey, the relational data can display how behaviours may be interpreted by some as inclusive experiences and as exclusive experiences by others. If patterns emerge in the subjective experience of certain demographic groups, otherwise covert inequality regimes can be rendered visible. This then allows organisations to tweak their diversity and inclusion interventions to adjust and balance inclusive and exclusive processes through a more nuanced data analysis than surveys alone may provide (Periac et al., 2018; Roberson et al., 2017). In practice, obtaining this type of behavioural dialogic data may prove difficult, and this is an ideal scenario description of how this type of data could be leveraged. If this type of data could be obtained though, it could help provide more nuance and sociological insight than routine survey data alone.

2.3: Gender-Based Violence

2.3.1: Definitions of Gender-Based Violence
Gender-based violence is a continuum, which includes everything from catcalling, to jokes perpetuating and normalising male domination, to unwanted touching, to “gray rape” (a controversial term for when someone does not or cannot say no), to forced sex (Kipnis, 2018, p. 198; 100 Women I Know, 20183). Gender-based violence is not always sexual in nature, but this research will mostly focus on this type of gender-based violence. University students have reported feeling confused on the differences between “sexual assault, rape, dating violence, domestic violence, and sexual abuse” (McMahon et al., 2021a, p. 3881). While rape is just one example of gender-based violence, the variability within the concept of rape highlights how gender-based violence as a whole is vast. For example within the EU, only 9 of 28 countries recognise consent-less sex as rape (Flecha et al., 2020). Scotland only de-gendered rape in 2009, while in Spain, rape is only considered rape if there is proof of physical force or intimidation (Flecha et al., 2020; Sanyal, 2019). As one female-identifying victim-survivor reflects, “I don’t think a man would think he’d raped someone unless it was physical and forced. So many guys would be horrified if they were made aware they had raped someone, I think it comes from a lack of education and understanding of consent” (100 Women I Know, 2018, p. 183). Both covert and overt violence cause harm, even if malintent is not present. To mitigate future harm, it is critical that interventions seeking to counter gender-based violence acknowledge the continuum and its more covert manifestations even with the best of intentions.

To begin to understand the gender-based violence spectrum and the intricacies of sexual consent, the literature highlights the importance of critiquing power and privilege hierarchies (Srinivasan, 2022; Hoxmeier et al., 2022; Flecha et al., 2020). Despite sex’s narrative of being a private act, it is still affected by the larger societal context (Olufemi, 2020). When it comes to binary genders, there are clear power hierarchies and gender roles; power does not magically disappear when it comes to sexual relationships (Burkett & Hamilton, 2012; Olufemi, 2020). Examples of this are ample, especially when it comes to falsely blaming victim-survivors. For instance, while women are expected to be desirable enough for the male gaze, men are the ones who grant women acknowledgement and attention (Given, 2020). This power hierarchy is seemingly reversed, however, when women are then blamed if they are assaulted because of dressing too desirably (Given, 2020). Model, author, and sexual assault victim-survivor, Emily Ratajkowski is infamous as one of the three naked dancing women in Robin Thicke’s Blurred Lines music video. After being sexually assaulted on

---

3 This book is a collection of anonymous accounts about gender-based violence written by victim-survivors. I have chosen to refer to the source by its title rather than the editors’ names. First, this is to emphasise just how widespread such experiences are. Secondly, I would like to hold the experiences and words of actual victim-survivors just as close as the academic literature in my writing.
the set of the video, she reflects, “Facing the reality of the dynamics at play would have meant admitting how limited my power really was—how limited any woman’s power is when she survives and even succeeds in the world as a thing to be looked at” (Ratajkowski, 2021, p. 47). Male superiority and sexual entitlement can be felt across gender, race, and sexuality, dictating who has power over whose body (Fiennes, 2019; Srinivasan, 2022).

Power hierarchies do not dictate, but they can influence sexual consent showing that there is nuance to the ‘one in four’ statistic well-known on university campuses (Mellins et al., 2017; Flecha et al., 2020). Power is going to inform who is allowed to revoke their consent (Srinivasan, 2022). Further, it informs who is most at risk (McCauley et al., 2019). People who are intellectually disabled are seven times more likely to be victim-survivors than people who are not (Levine & Meiners, 2020). Women and gender nonconforming students also report being victim-survivors at higher rates than men (Mellins et al., 2017). Coulter et al.’s (2017) study of college students found that 20.9% of transgender students were victim-survivors, while 6.4% of heterosexual students studied reported the same. White transgender students also had lower odds of sexual assault than Black transgender students (Coulter et al., 2017). Re-victimisation is also high for university students across gender, with one study of university students reporting the median number of sexual assault instances as three independent experiences of assault (Mellins et al., 2017). Such differences in power, and one’s resulting risk of violence and re-victimisation, are directly reminiscent of inequality regimes; a university campus is going to have a unique inequality regime that influences gender-based violence.

2.3.2: Theoretical Grounding of Bystander Interventions for Gender-Based Violence

How bystander trainings are introduced and framed to participants seems to have a direct impact on how effective the trainings are. One study found that if the bystander training frames itself as a prevention, rather than as a response, the training will have more of an impact on participants (Alegría-Flores et al., 2017). Similarly, researchers found that when conducting the training with students, it is worthwhile to anchor the training in leadership and feminist frameworks to connect the content to students’ interpersonal skills (Eriksen, 2021). For students who are less motivated by the social justice or interpersonal reasoning to intervene, it may also be helpful to connect the bystander training to the concept of morality (Brodt et al., 2021). This is especially impactful for more socially conservative students, who may come from more reserved cultures or religions (Brodt et al., 2021). This may involve highlighting how the potential harm relates to the fairness of a situation, or the sanctity of a possible victim-survivor (Brodt et al., 2021). Introducing the concept of blame provokes students to
shift the idea of who is at blame: minimising the blame of the victim-survivor, and shifting the blame to that of the perpetrator and themselves as an inactive bystander (Holtzman, 2021). A shared responsibility can be fostered by asking participants to make a commitment to becoming active bystanders, thus promoting long-term behavioural changes (Banyard et al., 2004). This approach, like the leadership, feminist, and morality frameworks too, help create a more community-oriented approach to being a bystander; urging participants to consider how if they do not intervene, they then are acting as an accomplice in the assault (Holtzman, 2021). Unfortunately, this community approach is overlooked in much of the gender-based violence and bystander literature, which instead favours a five-step framework to understand how a bystander decides to intervene (Banyard, 2011).

One of the leading frameworks related to this idea of community though is Banyard’s (2011) ecological model of bystander intervention. This framework, modelled on Bronfenbrenner (1977), will guide this research as it emphasises not just the individual-level tactics necessary to address structures that hold gender-based violence in place, but also the community-oriented approach (Amar et al., 2015; Banyard, 2011; Bonar et al., 2022; Klein et al., 2022; Hirsch & Khan, 2020). The ecological model organises bystander behaviour into three levels; intrapersonal factors within a person, microsystem factors which focus on relationships, and exosystem factors which look at the larger community setting (Banyard, 2011). Incorporating this community-oriented model is critical as Banyard et al. (2021) found that perceived peer helping and perceptions of community influence are correlated with bystander prosocial actions in scenarios with just friends and just strangers alike. Gainsbury et al. (2020) also researched how focusing on general community members can help prevent sexual violence across diverse community contexts, while some researchers call for a particular focus on nightlife settings like bars (Bonar et al., 2022; McMahon et al., 2021b). The importance of the exosystem is highlighted by the fact that bystanders are present in only 17% of sexual violence cases, while university students are more likely to be victim-survivors of sexual violence beyond the perimeters of their campuses (Bonar et al., 2022). As a result, while individuals must feel prepared to act on behalf of themselves, universities should target not just student organisations, but all students and community members on and off-campus to address social norms (Banyard et al., 2004).

With the ecological model of bystander intervention, there is more potential to investigate microsystem and exosystem factors, such as collective efficacy and trust, that influence gender-based violence prevention too. At the microsystem level, collective efficacy, which encompasses trust of neighbours, was significantly associated with bystanders being helpful in certain scenarios (Banyard, 2016). Sulkowski (2011) and Sapouna (2010) also both found that trust between community
members reduces violence and also promotes helping behaviours. In the military context, Holland et al. (2016) found that rank and unit morale are predictors for bystander intervention too. At the exosystem level, greater trust in the institution itself and law enforcement is also a predictor for helpful bystander intervention (Holland et al., 2016; Sulkowski, 2011). Yet, one Georgetown study found that trust in the university and some support services actually decreased after a bystander intervention training program (Smith and Powers, n.d.). Together, these studies show a necessity for universities to investigate microsystem and exosystem factors as they are highly under-researched, but do influence how rampant gender-based violence may be (Khan et al., 2020).

Relating this back to measurement scales of behaviours and perceptions, it is clear that higher education institutions tend to focus only on individual-level behaviours and perceptions given the lack of a widespread community-level framework (Khan et al., 2020). Additionally, it is common practice and to an extent mandated in the U.S. for HEIs to keep track of statistics like reported rapes and administer campus climate surveys students to understand general knowledge, awareness, and perceptions of their behaviours (Tilley et al., 2020; Woods et al., 2022). On the other end of this, UK institutions are plagued by a lack of reporting (Donaldson et al., 2018). In this way, there is an opportunity for UK institutions to explore student perceptions and monitor actual behaviours too, especially when it comes to bystander behaviours. Jerolmack & Khan (2014) found that of 140 studies of sexual assault prevention programs, just 17 studied participants’ behaviours instead of attitudinal measures. Furthermore, just as in inclusion research, there is room in gender-based violence research to ensure behaviour measurements are integrated in reporting mechanisms from the get-go. In doing this, gender-based violence research can move beyond individually-focused reporting measures to ensure the wider social context is taken into account.

2.3.3: Sexual Consent & Active Bystander Interventions

Universities’ options for sexual consent and bystander intervention training are endless. The overarching sentiment towards the number of training options is summarised in the statement by one U.S. university’s Title IX coordinator: “I wish that companies would stop hawking their ‘trainings’ and we could all agree upon something that is really good.” (Chambers et al., 2021, p. 37). There is much decision-making that goes into universities choosing a programme. They are not just looking to protect students, but also to protect themselves in choosing a programme: legally, in terms of reputation, and financially as bureaucratic institutions (Srinivasan, 2022). Popular off-the-shelf programmes most often cited in research studies (mostly from the US & UK) include: Bring in the Bystander, The Green Dot Program, The Women’s Program, The
Men’s Program, One Act, Helping Advocates for Violence Ending Now, Friends Helping Friends, SCREAM Theatre, Mentors in Violence Prevention, Step UP!, Speak UP!, Coaching Boys into Men, Campus Clarity, InterACT, RealConsent, Consent and Respect, Stand Up and Speak Out, and Take Care, among many others (Evans et al., 2019; Amar et al., 2015; McMahon et al., 2021a; Foubert et al., 2010; Eriksen, 2021; Stojanov et al., 2021; Alegría-Flores et al., 2017; Mujal et al., 2021; Acquaviva et al., 2022; Katz & Moore, 2013). Just as with inclusion programme interventions, many of these bystander programmes are not informed by theory (Evans et al., 2019). One meta-analysis found that just 56.8% have an underlying theoretical framework (Mujal et al., 2021).

There is a desire for more evidence-based interventions in both the literature, and in the field as evidence for intervention effectiveness is mixed (Bonar et al., 2022; Htun et al., 2022). Bystander trainings are overall effective, but of course, this is dependent on how one defines effectiveness. Studies most often explore effectiveness in terms of increasing knowledge, understanding, confidence, pro-bystander behaviours and attitudes, and decreasing rates of victimisation. Some examples of the most common measures used to evaluate programmes are: the Bystander Behaviour/Knowledge/Attitudes Scales; the Bystander Self-Efficacy/Confidence and Intent/Willingness to Help Scales; the Illinois Rape Myth Acceptance Scale; the Perceptions of Peer Helping Scales, and the Modified and Original Readiness to Change/Help Scales (Mujal et al., 2021; Sulley et al., 2020). Just as with the inclusion measurements, these scales often over-rely on student and administrators’ perceptions. However, even if trainings are deemed effective overall by one of these subjective scales, the extent to which this is achieved is a mixed bag and found to be modest at best (Htun et al., 2022).

While bystander trainings are effective to varying degrees, it is important to note that confidence and intentions do not always translate into long-term, or short-term for that matter, behavioural changes (Katz & Moore, 2013; Wong et al., 2021). Jerolmack & Khan (2014, p. 200) refer to this as “attitudinal fallacy” whereby reports of someone’s sentiments are sometimes mistaken by researchers as directly indicative of one’s behaviours in given situation. In the short term, bystander training participants report increased confidence across the board (Evans et al., 2019). The extent to which confidence increases, and for how long, is highly varied though (Evans et al., 2019). This remains true too when it comes to long-term effect on behaviour, efficacy, and willingness to intervene (Evans et al., 2019; Mujal et al., 2021). One study focusing on The Men’s Program found that two years after training, 20% of participants felt no attitude or behavioural changes (Foubert et al., 2010). A more recent study found that training participants experience an increased awareness of alarming behaviours, desires to help, actual helping and pro-active learning behaviours, and changes in attitude and knowledge (McMahon et al., 2021b). Another study evaluating the Bring in
the Bystander programme found an increase in knowledge, understanding and efficacy, but no real change in behaviours, further hinting at bystander interventions’ mixed results in the literature (Stojanov et al., 2021). One of the most promising long-term results comes from one multi-year multi-campus study in which a campus that trained just about 15% of its students had 17% lower rates of victimisation after a four-year period compared to the control campus (Coker et al., 2016). Consequently, the literature repeatedly finds that trainings do work, but the extent of this is highly variable.

Furthermore, a few distinct barriers to intervention emerged in the literature that remain even post-training. While these obstacles to intervention should be directly addressed by universities in their sexual consent and bystander training programmes, they often are not. Kania & Cale (2021) and Blayney et al. (2021) both have found that barriers to intervention include failing to notice, being ignorant of gender-based violence markers, ambiguity of the situation, failure to take responsibility, a skills deficit, and failure to intervene due to an audience. To add to this, intoxication and its effect on ambiguity reading also impede intervention (Blayney et al., 2021). This struggle of “being sure” and hesitant when it comes to alcohol and consent was highlighted in other studies too (Brodt et al., 2021, p. 13; Foubert et al., 2010; Hillman, 2021; Walsh et al., 2021). In the US, 50% of university sexual assaults involve the consumption of alcohol, while consumption of 1 unit of alcohol in men is correlated with 13% higher odds of sexual violence perpetration (Hillman, 2021; Steele et al., 2022). Each of these critical barriers to intervening can and should be directly discussed and triaged in university modules. In doing so, students will be better equipped with the skills to act as future active bystanders in both straightforward and trickier situations.

The limitations of sexual consent and bystander interventions and their resulting measurement scales is very similar to the limitations of the inclusion research: the literature fails to investigate the larger context and measurements remain too individually-focused as a result. To move the gender-based violence literature forward, researchers could incorporate the ecological model of bystander intervention to pursue more community-oriented factors that may influence bystander intervention. As previously mentioned, these could pertain to the microsystem (ex. relationships) or the exosystem (ex. the institution, the city, etc.). In overlooking these social realities, bystander intervention research outright neglects that gender-based violence is relational and contextual. Furthermore, measures used to evaluate bystander interventions are overwhelmingly subjective, given that most training evaluation methods rely on participants’ self-reporting (Evans et al., 2019). Just as with inclusion research, there is a glaring opportunity for researchers to expand measurements beyond that of individual perceptions of intervention effectiveness. Bringing in behavioural measurements is key to moving the gender-based violence research forward.
2.4: Moving Theory Forward with Mixed-Methods

In reviewing the literature, it appears there is space for a new way of investigating gender-based violence and inclusion/exclusion while being guided by relational inequality theory (Tomaskovic-Devey & Avent-Holt, 2019). The framework I propose here studies gender-based violence and inclusion/exclusion in terms of covert vs. overt, and individual-level vs. group-level. In doing so, it expands the measurements possible to include group-level behavioural measurements. Some theories have begun to step towards this new theoretical framework. With the legitimacy theory of inclusion/exclusion, Van Dijk & Khattab (2021) approach inclusion/exclusion in terms of behaviours and perceptions. This inches closer to an investigation of both covert and overt displays of inclusion/exclusion. Banyard’s (2011) ecological model of bystander intervention introduces group-level understanding of gender-based violence, helping the literature move beyond the limitations of an individual-level only framework. Keeping these two theories and relational inequality theory at the forefront of my work further exhibits how a portion of investigations into social closure, namely opportunity hoarding and exclusion, can be easily missed by researchers. Conversely, researchers instead tend to focus on overt and individual-level manifestations of claims-making, exploitation, and social closure. Notably, at the more individual-level, certain behaviours are easier to see face-to-face in interactions, while at the group-level, even if overt, manifestations can become more institutional or systemic. This is not to say claims-making or exploitation do not occur covertly at the group-level. Rather, the manifestations of social closure at the covert group-level present an under-researched area that we as researchers can course correct with computational social science. I will expand on this with concrete examples of both social phenomena, but this theoretical framework may be visualised as four quadrants as in Figure 2.1.
Figure 2.1: A New Theoretical Framework of Inclusion/Exclusion and Gender-Based Violence

To highlight the forces behind group-level overt manifestations of inclusion/exclusion requires starting with common examples, as seen in Figure 2.2. In the university context, group-level overt exclusion and opportunity hoarding may be certain demographic groups being rejected from the university at increased levels in comparison to others. An example is students from low-income backgrounds or racial minorities may be less represented in elite universities. These are institutional items that can be easily monitored by university administration. Social closure does not solely exist in the overt and group-level though, another example can also be seen in the overt and individual-level when a tutor spotlights a student in class based on their identity. This would be including the student in conversation, but excluding them from the majority group and inadvertently labelling them as ‘different’ or ‘other’. It could also be exploitation, by exploiting the student’s identity, without consent, for others’ gain in knowledge. A more concrete example of this could be if in a criminology class, a teacher asks the only Black student, “What do you think about George Floyd’s death and police violence?”. Some students may appreciate being spotlighted, while others can interpret it as very exclusionary. It is important to differentiate here that this is an individual-level manifestation of exclusion because it is a behaviour directed at one person within an interaction, even though that behaviour is based on the individual’s larger group identity.

Moving along to the individual-level and covert manifestations also shows how these forces are at play in the university context. Individual-level more covert exclusion may be mispronouncing someone’s name or not using their correct pronouns. Again, it depends on the individual, but over time, behaviours like this may contribute to a
person’s feelings of exclusion. It could also be seen as claims-making in a way, especially if the person refusing to respect the pronouns or mispronouncing the name is from a majority demographic group. The person from the majority would be laying claim that their knowledge of that person’s identity is superior to the person who is being identified. Microaggressions are an additional well-known form of covert racism gaining traction in the academic literature, which comprise brief slights that have insulting undertones based on someone’s demographic group (Garcia & Johnston-Guerrero, 2016). Therefore, microaggressions exclude others by inadvertently labelling them as ‘other’ through the mechanism of social closure. This could include a teacher saying “You don’t look disabled” to a student who informs the instructor of, for example, their non-visible chronic illness or mental health status.

When the new theoretical framework is applied to inclusion/exclusion, covert group-level social closure is pinpointed as understudied. A common example could be systemic exclusion such as a student not being spoken to in their university class. This could be exclusionary as students may just be labelled as an introvert, and for that reason, not be pulled into classroom discussion. At the same time, it is also opportunity hoarding as most of the knowledge production is being claimed by other students. If this is occurring along racial and ethnic, gender, disability, or any other salient demographic line at the group-level, then it is important to address. Yet, the sheer scale of this is more difficult to see and monitor with current inclusion/exclusion measurements. Furthermore, while this example is rooted in a face-to-face behavioural example, if it is happening at a grand scale across university classrooms, it is more difficult to pinpoint as it becomes normalised. This is ultimately why covert and group-level exclusion becomes significantly harder to both study and address because it shifts towards becoming systemic.
Similarly in Figure 2.3, applying the theoretical framework to gender-based violence highlights how insights into social closure at the covert group-level of gender-based violence is limited in the literature. Again, this is not to say claims-making and exploitation do not happen at the group-level covertly. However, arguably, the most studied examples of gender-based violence are the more overt claims-making displays at the group-level and individual-level. For example, at the group-level, the number of gender-based hate crimes or reported rapes on campus is a statistic some universities would have on file. These figures collectively showcase the (reported) amount of claims-making happening on campus. Overt, but individual gender-based violence indicators could be things like stalking, verbal violence, rape, as well as the more controversial “gray” rape. These manifestations of claims-making are easier to see, note, and address at the face-to-face level. At the same time, while rape is widely agreed to be gender-based violence, one person may consider an interaction to not have been rape, while another may have felt violated. This type of rape may still be reported, but it is common for the classification of the experience to be dependent on a person’s own perspective of the interaction (Khan et al., 2020). More covert individual-level examples involve social closure, as well as claims-making. These include microaggressions and gender-based violence “jokes”, which may be likened to exclusion and claims-making. The perpetrator in this case would most likely be deeming their counterpart ‘other’ and inferior in some way, while claiming their own social position to be the right or better position to be in. While the behaviour is critical to
monitor, an individual’s own perspective is important to determine whether gender-based violence is occurring in these more interactional cases.

Finally, more covert group-level experiences can be as simple as the presence, or lack thereof, of trust. Trust, in this case, relates to one’s confidence in an entity or another person behaving in the way that is expected (Mohammadi & Hashemi Golpayegani, 2021). Two relevant types of trust to this quadrant of the theoretical framework include social trust and institutional trust. In the case of gender-based violence within higher education, social trust is one’s trust in others, likely a peer group or general student population. Institutional trust is a student’s trust in their higher education institution to protect students from, and handle, incidences of gender-based violence. While institutional trust dictates how secure a student feels towards their higher education institution, it also influences the social norms through which a student makes sense of their peer relationships (Banyard, 2011). In this way, the exosystem influences the microsystem (Banyard, 2011). For example, if a student perceives their institution as trustworthy due to their gender-based violence strategy (including a mandatory bystander intervention training), the same student may feel more secure in their peer’s ability to recognise and mitigate gender-based violence too. Conversely, if students are distrusting towards their institution, perhaps they then are less inclined to engage with gender-based violence programming if it is even available. Both scenarios would both covertly and directly impact the culture of and success of any potential bystander intervention.

This takes Banyard’s (2011) ecological model of bystander intervention a step further by accentuating how covert factors influencing the group-level can be sometimes. The covertness of trust, whether social or institutional, is accentuated by both mechanisms of social closure in this case. On one hand, students who feel less trust may be excluded from the social circles or networks that feel most trustworthy and safe. On the other hand, the more trusting peer groups or institutional networks may then be reserving any potential protection from gender-based violence for their members only. Given the covertness of these social dynamics and experiences, researchers often overlook this and the impact they have on incidences of gender-based violence. Only recently have researchers begun to study this bottom left quadrant where the covert and group-level overlap to create gender-based violence related social closure. Critically, while research is needed on overt behavioural gender-based violence, there is also an evident need for research into the harmful consequences of distrust as it relates to gender-based violence.
Computational social science has gained traction in sociology since 2010, and I seek to show how incorporating these approaches into mixed-methods research will help investigate the group-level covert behavioural manifestations of these social phenomena (Edelmann et al., 2020). Although gaining traction, there is a lack of computational training in sociology curricula as a whole (Keuschnigg et al., 2018). Computational social science is inherently interdisciplinary. The simplest explanation of computational social science is it is a combination of social science theories with methods borrowed from computer science (Edelmann et al., 2020). These techniques are usually applied to very large datasets (labelled ‘big data’), but not always (Edelmann et al., 2020). There is a consensus in academia that computational social science techniques offer a promising advancement of sociology (Keuschnigg et al., 2018).

Two theoretical frameworks highlight the potential of computational social science for the studying of sociology with mixed-methods: relational sociology and analytical sociology. Both theories explore macro-level phenomena by first studying the micro-level. Relational sociology’s focus is mainly on the interactions between individual actors, while analytical sociology researches the individual actors and their interactions (Crossley, 2016; Keuschnigg et al., 2018). The latter also places more emphasis on how the micro-level dynamics can reveal more unobservable collective dynamics (Keuschnigg et al., 2018). In much the same way that ethnographies observe experiences and processes directly, computational social science methodologies can be used to examine underlying and overlooked behaviours of a social phenomenon (Keuschnigg et al., 2018). Tools to do this were limited until computational social
science, as surveys are limited in their ability to show the evolution of perceptions and behaviours without creating survey fatigue (Edelmann et al., 2020; Keuschnigg et al., 2018). Drawing on my previous arguments, this is not to say surveys are unnecessary. Rather, a mixed-methods approach following the logic of relational and analytical sociology allows for a computational extension of qualitative methods whereby researchers can more effectively reveal otherwise ‘hidden’ inequality regimes. Within a given network, differences in power, resources, and other forms of capital can come to light with mixed or “hybrid” methods that incorporate computational social science techniques (Crossley, 2016; Edelmann et al., 2020, p. 73). Therefore, computational social science can help researchers study micro-level behaviours and interactions, which can help drive sociological theory at the macro-level (Flache et al., 2022).

2.4.1: Social Network Analysis and Inclusion

Social network analysis (SNA) is the empirical study of networks and offers researchers a means to relate micro-level behaviours to macro-level phenomena such as inclusion (Crossley, 2016). Embedding relational sociology in social network analysis requires the focal point of study to be the interactions between individuals or groups of actors (Crossley, 2016). The micro-level behaviours and interactions, or lack thereof, between individuals are hence the points of interest (Crossley, 2016). As discussed previously in this chapter, if inclusion researchers anchor inclusion in perceptions of uniqueness and belongingness, as well as contribution and participation, traditional survey methods to measure inclusion can be coupled with aggregate and more reproducible computational approaches such as social network analysis. In doing so, inclusion researchers not only overcome the lack of reproducibility of surveys, but also utilise social network analysis to promote the legitimacy theory of inclusion. This mixed-method therefore helps promote the studying of micro-level behaviours to determine the legitimacy of inclusion and exclusion of entire demographic groups.

Given my new theoretical framework pushes for the studying of behavioural manifestations of inclusion/exclusion, I also call for the use of SNA to overcome the overemphasis on perceptions of inclusion/exclusion to date. SNA allows researchers the option for the unit of study to become relational data between individuals (Wagner & González-Howard, 2018). Interactions as key parts of behaviours help construct the social context of the classroom (Wagner & González-Howard, 2018). With each class, and the unique set of attendees, patterns of interaction will be reconstructed and shift each session depending on who shows up to class. Inclusion and exclusion, which comprise otherwise unobservable or covert divisions between individuals become more observable with SNA (Akemu & Abdelnour, 2020; Crossley, 2011). Furthermore, SNA as a methodology can help identify which critical individuals may be key in the promotion
of more equitable access to resources and opportunities, therefore who is key in mitigating opportunity hoarding (Shantz et al., 2011).

Despite the promise of SNA, using SNA to study inclusion remains uncommon. Yet, borrowing SNA from the data science field fosters sociologists’ ability to identify how micro-level behaviours and relations can influence macro-level systems of hierarchy in a given inequality regime (Jacobs & Watts, 2021; Page, 2015). Even with SNA’s reproducibility and applicability to inclusion research, SNA has been sparingly used to measure inclusion within the education and university space, where diversity and inclusion literature originates (Chrobot-Mason et al., 2013; Ferdman, 2013). One study by Karimi & Matous (2018) used social network analysis to visualise the interactions of university students in their extracurricular activities, and in doing so, found evidence of homophily between international students who tended to interact more with their co-nationals than home students. Other noteworthy studies by Collins & Steffen-Fluhr (2019) and by Hardcastle et al. (2019) attempted to assess the behavioural changes of collaboration networks before and after interventions at universities, only to find that female researchers were still less often chosen as co-collaborators after human resource interventions. These studies are just some examples of how social network analysis can be leveraged to identify patterns of inclusion, with the latter studies distinctly calling out “[…] disturbing gender patterns that were previously invisible and difficult to quantify” (Collins & Steffen-Fluhr, 2019, p. 275; Tienda, 2013).

Rarer is the usage of SNA to study real-world learning settings through dialogic patterns, capitalising on both the behaviours of and perceptions of inclusion in the classroom (Wagner & González-Howard, 2018; Zhang et al., 2022). In fact, the only study to have used a mixed-methods approach of SNA in combination with a survey to understand students’ experiences in an online forum setting failed to also incorporate demographic analysis (Dawson, 2008). Demographics influence perceptions of inclusion in professional networks, so there is a case for researching how this translates to classroom networks too (Jung & Welch, 2022). Furthermore, SNA, when applied to real-world classrooms, mainly focuses on teachers rather than the student experience (see Frøytlog, 2022; Mameli et al., 2015, Pantić et al., 2021; Zhang et al., 2022). While Mameli et al. (2015) looked at nodal degree, degree centrality, information centrality, degree centralisation, and cliques to find that teachers act as conductors of student interactions, the study does not incorporate the students’ subjective experiences of their interactions.

As discussed above, there is a gross underutilisation of methods that allow for the study of group-level and behavioural manifestations of inclusion/exclusion. At this time, surveys remain too individually-focused on perceptions. Instead, a mixed-methods approach with SNA is a more reproducible method that can be used to study more
covert behaviours, as well as compare these behaviours of inclusion to perceptions of inclusion (Fuhse, 2023). My inclusion study’s ultimate intent is to do just that by collecting both surveys and dialogic data for social network analysis to study and compare patterns of inclusion after a training intervention. Surveys still can and should continue to be used to study the perceptions of inclusion within the classroom to add much needed context to the SNA data. However, only with the addition of social network analysis of inclusion will I be able to cross-compare if and how covert behaviours of inclusion may differ from one’s perceptions of inclusion. Furthermore, keeping students and teachers, along with their demographics, at the forefront of both of these methods can help researchers better understand how covert inequality regimes may or may not be adopted into the classroom setting within higher education institutions. Thus, there is an outstanding need for the inclusion literature to incorporate SNA into real-world learning settings and ensure a mixed-methods approach that looks at both the behaviours of and perceptions of group-level inclusion.

2.4.2: Computational Text Analysis and Trust

Computational text analysis presents a new way of finding meaning in the written word, whether that written word was originally a text or a transcription of dialogue. Incorporating this technique in sociology is uncommon when compared to other more qualitative text analysis approaches (Schwemmer & Wieczorek, 2020). If gender-based violence research expands to incorporate the ecological model of bystander training, there is potential for researchers to begin to utilise computational text analysis methods. This offers a more reproducible method than qualitative approaches and can be applied to larger datasets without significantly increasing analysis time. As mentioned previously, the ecological model of bystander intervention looks at more community-level factors, such as trust. Velner et al. (2021) explores three measures of trust: subjective, active objective, and passive objective. Subjective relies on self-reporting of trust, active objective explores actions taken between individuals, and passive objective investigates dialogue between actors (Velner et al., 2021). Expanding to the ecological model means that trust within dialogue at a university can be researched to understand why gender-based violence may be persisting. Incorporating my new theoretical framework also refocuses the research on group-level and more covert manifestations of trust. Afterall, trust is relational and does exist at the group-level, even if someone is not overtly discussing it (Mohammadi & Hashemi Golpayegani, 2021). For instance, there is an opportunity to research whether students distrust other students or their institution as a whole, or if there is an overwhelming trust in others and their
Given CTA's usefulness in studying difficult to assess social phenomena like trust, CTA has been used to study trust before. However, research to date lacks dialogic application, is mostly restricted to digital communication data, and has not been applied to trust within the gender-based violence context. Interestingly, most trust-based computational text analysis work has been applied in the e-commerce space to understand consumer trust. Outside of that, Mohammadi & Hashemi Golpayegani (2021) used sentiment analysis to train a model to measure trust between Reddit users in one online network. Alsaid et al. (2022), who I draw significant inspiration from, created a lexicon of trust-related words using word embeddings from 46 trust questionnaires that measured human-human trust, e-commerce trust, and trust in automation. The researchers call for an application of their trust-related lexicon to measure trust in conversations between people. Li et al. (2020) take a step towards this in their paper on measuring trust between an AI customer service chat and a human user. Yet, to date, there remains a gap in the literature in terms of computational text analysis to measure trust in human-to-human conversations. This is despite robust processes in place to develop unique trust-related lexicons dependent on the type of data, whether e-commerce, human-to-human or automation-related.

With that, I find that there is a gap in the literature in applying CTA to measure trust within dialogue about gender-based violence. Keeping in mind my theoretical framework, there is an opportunity to expand CTA to the covert group-level forces of gender-based violence. This could be done by applying dictionary-based trust lexicons like the NRC dictionary to group conversations about trust as it relates to gender-based violence. If a pre-made trust dictionary is not used, a bespoke dictionary could also be developed in a similar manner to Alsaid et al. (2022). Indeed, this is what my gender-based violence study intends to do to; ultimately, to either reinforce or challenge complementary qualitative thematic analysis. By specifically pairing thematic analysis with computational text analysis in this study, my work acts as a potential proof of concept where trust in the context of gender-based violence may be investigated solely with CTA in the future if proven to complement my qualitative findings as ground truth. In instances where qualitative analysis may be too time consuming, CTA therefore could offer an alternative reproducible method. The specific combination of CTA with thematic analysis in this case though also could help add tangible figures to qualitative findings by providing an estimate of trust pre- and post-training intervention. Additionally, while the two methods allow insight into students’ trust in the general student body and institution, it is worth acknowledging that there are aspects to gender-based violence and the inequality regime that will not be
captured. This includes the number of campus assaults, the actual incidences of bystander intervention on campus, and the trust students maintain within their specific friend groups or their student body organisations.

To achieve the above, gender-based violence related documents discussing trust can be used to create a gender-based violence trust dictionary. Then, employing sentiment analysis to dialogue between students about gender-based violence provides insight into how much trust students feel within their given higher education community. This will be done with the same group before and after a gender-based violence training intervention to understand how levels of trust shift for students within a higher education institution. As they learn new bystander skills and sexual consent knowledge, this may affect how much trust they feel in other students at their institution, as well as how much trust they feel towards their institution in protecting the student body from gender-based violence. By combining CTA and thematic analysis, institutional trust and social trust pre- and post-training can thus be investigated through a robust two-prong approach.

### 2.5: Chapter Conclusion

This chapter builds on Acker’s (2006) inequality regimes, Tomaskovic-Devey & Avent-Holt’s (2019) relational inequality theory, and Bourdieu’s symbolic violence (1998) to explore why gender-based violence and inclusion/exclusion are critical for sociologists to study. Both social phenomena are influenced by an institution’s inequality regime which dictates hierarchies of power, oppression, and domination. While I acknowledge current frameworks for inclusion and gender-based violence, I emphasise that two frameworks in particular will help move the literature forward. Namely, van Dijk & Khattab’s (2021) legitimacy theory of inclusion and Banyard’s (2011) ecological framework for bystander intervention. I highlight how these two frameworks, if adopted by researchers, offer a more nuanced understanding of both concepts. Van Dijk and Khattab’s (2021) legitimacy theory continues to anchor inclusion/exclusion in an individual’s experience of the processes, but also calls for the measurement of behaviours of inclusion/exclusion. Similarly, Banyard’s (2011) ecological framework exposes how bystander interventions must take into account the larger community context. As a result, research would look beyond any one individual’s interpretation. After discussing how these two frameworks move the literature forward, I propose my own new theoretical framework for gender-based violence and inclusion/exclusion. I show how to further move the literature forward, inclusion/exclusion and gender-based violence researchers must incorporate more inquiries into the more covert and group-level manifestations of these social phenomena. I show how explorations into the
covert and group-level would help uncover mechanisms of social closure specifically for both social phenomena, while also accentuating research into both behaviours and perceptions.

In providing my theoretical foundation, I explore how new measurements can be applied to inclusion/exclusion and gender-based violence research to expose inequality regimes before and after interventions. Specifically, I share how my new theoretical framework allows for researchers to integrate computational social science techniques into our field through mixed-methods. Namely, researchers must and can study behaviours of inclusion/exclusion by utilising social network analysis. Similarly, computational text analysis offers a way to measure covert and group-level factors, such as trust within dialogue, that can influence whether gender-based violence occurs or not. With these two computational methodologies added to my mixed-methods research design, I provide the groundwork for answering my research questions while emphasising both objective behaviours and perceptions.

Chapter 3: Methodology

This chapter will extensively discuss my methodological decisions given the methodological focus of my research. In this chapter I will address the separate, but related, data collection processes for both the inclusion study and gender-based violence study. Each study was conducted at the same higher education institution. Across the two studies, I collected data through semi-structured focus groups, individual semi-structured interviews, surveys, field recordings, and ethnographic observations. I utilise mixed-methods analysis not just due to the varied data collected, but also due to the research questions’ exploratory attempt to highlight how computational social science may be integrated into inclusion and gender-based violence research. Eclectic by nature, my work uses thematic analysis, statistical analysis, as well as the computational methods of social network analysis and computational text analysis. Finally, this chapter provides space for honest reflection on critical ethical considerations.

3.1: Field Site Setting

Given the sensitive nature of the topics of gender-based violence and inclusion/exclusion, I anonymised the study’s participants by default. While anonymisation is critical in protecting the wellbeing of individual participants, my decision to protect the anonymity of the institution is not taken lightly. I align myself with Baez (2002) in that I seek to protect the vulnerability of the individuals who graciously donated their time in helping me expose power structures within academia. I
deeply regret not being able to render the institution itself vulnerable enough to encourage accountability and transformation. Ultimately, anonymisation of the field site is a necessity in protecting the identities of the study’s participants, especially crucial for those who at the moment of writing are still actively navigating historically oppressive institutions as historically marginalised individuals. Unfortunately, this directly limits this study’s ability to adopt a truly transformative research design by not creating maximum accountability for the university (Baez, 2002).

It is still important to share some high-level attributes of the participating university. The participating institution is an elite public higher education institution in the United Kingdom. As part of the Russell Group, the university prides itself on the usual interests of most historically white, high-ranking, and well-funded 21st century educational institutions: being research-driven, global, future-oriented, and a leader in higher education and learning. Coupling the university’s desire to be a leader in higher education with the fact that inclusion and gender-based violence are ‘hot topics’ within academia, it is unsurprising that the partner university both sought out and approved this intervention research. Please see Appendix A.1 for a short review of the other pre-existing strategies at the university for mitigating gender-based violence and fostering inclusion.

3.2: Gender-Based Violence Study Research Design

Data collection occurred with two cohorts of students and included semi-structured focus groups, individual semi-structured interviews, and surveys. The initial research design was chosen internally by a university committee to evaluate the effects of the Epigeum Consent Matters & Tackling Harassment training intervention, but I adapted this research design with input from the committee. The core staff committee consisted of the head of the school, a counselling representative, and a student careers representative involved with other student trainings. The committee’s proposed research design initially included only focus groups and surveys, the latter of which is the norm in many gender-based violence evaluation studies (Sulley et al., 2020). The 1:1 individual interviews were later added after data collection began for Cohort 1. This was due to the low participation rates in focus groups for Cohort 1. Due to the higher number of participants and less time available, no 1:1 individual interviews happened for Cohort 2. For an overview of the research design per cohort, please see Appendix B.1. The primary researcher for the study was myself as the PhD researcher. The university funded two different research assistants to assist at different times throughout the 1.25 year study.

The online self-paced training itself consisted of two parts: Consent Matters and Tackling Harassment. Consent Matters focused on teaching sexual consent,
boundaries, respect, communication, and introduced bystander intervention. It had three subsections: ‘Thinking About Consent’, ‘Communication Skills and Relationships’, and ‘Looking Out for Others’. Tackling Harassment further aimed to develop students’ bystander intervention skills by teaching students how to recognise problematic situations, to overcome common barriers, and to respond to disclosures. The two subsections for this module were ‘Being an Active Bystander’ and ‘Responding to Disclosures’. Both sections of the training incorporated scenario-based activities, educational materials in the form of further online reading, and lecture-style videos. The training was instituted in the hopes of mitigating gender-based violence and creating a values-based university community focused on inclusion, trust, and safety. The training was implemented in August 2021, and the research design process began in November 2021. The training remained voluntary for all staff and students, but the research team utilised these research findings in the consideration of whether to mandate the training in the future.

3.2.1: Participants

The sampling method chosen was purposive sampling, as the research team had clear characteristics we wanted to recruit for. Namely, we wanted first year students who lived in student accommodations. This allowed the research team to utilise the partnership with university student accommodation to boost participation quickly, as well as pinpoint perceptions of gender-based violence within the university community directly. First year students were the student population targeted given this demographic makes up the majority of student accommodations and would be the target audience should the training be mandated in the future. Student accommodation halls are off-campus buildings completely made up of students. It was also assumed that first years would feel more enthusiastic towards participating in an optional university training. For Cohort 1, only six first year student accommodations were initially recruited from. This was to ensure a diverse group of accommodations, especially considering that student accommodations at the university varied widely in price point (£4,333–£10,267 annually), catering styles, and location. This was also with the assumption that students would be keen to participate given the campus climate of gender-based violence discussions within the previous academic year. Recruitment fell short of expectations though despite incentives, meaning that the study then expanded its scope to all first years living in student accommodation. We instead recruited from 36 distinct student accommodation halls. A copy of the consent forms utilised may be found in the Appendix B.2 and for more information about study communications and incentives, please see Appendix B.3.
Participants were very diverse across several dimensions of diversity. While demographics were taken at the beginning of the study, and post-training, the demographics discussed here are from the start of the study. Two students were removed from the study post-study completion as it was revealed they were visiting students from other universities. This brought the total number of participants across both studies to 32 students. Student age ranged from 18 to 39. 53.13% of students were 18 years old at the start of the study, but the mean age was 19.78. 62.50% of students identified as female, 25.00% male, 9.38% non-binary, and 3.13% transgender. These gender demographics are relatively aligned with official university reports, as the undergraduate population is reported to be 61.55% female and 38.18% male. 43.75% of students identified as heterosexual, 28.13% bisexual, 12.50% queer, 6.25% asexual, 6.25% gay, and 3.13% pansexual. 68.75% of students reported being single, while 21.88% reported being not single. Of those, students reported casually dating one person (n= 3), casually dating multiple people (n= 2), and being monogamous (n= 4). One student did not respond. Of the students who were not single, the average relationship duration was 16.56 months. Students were asked to provide the length of their longest relationship if they were actively dating multiple people such as in polyamorous or non-exclusive relationships.

The study asked students to self-identify as: White, White-Scottish, White-Other British, Black or Black British- Caribbean, Black or Black British- African, Other Black Background (option to fill in), Asian or Asian British- Indian, Asian or Asian British-Pakistani, Asian or Asian British- Bangladeshi, Chinese, Other Asian Background (option to fill in), Mixed- White and Black Caribbean, Mixed- White and Black African, Mixed-White and Asian, Latinx, Arab, Gypsy or Traveller, Unknown, Information Refused, or Other Racial & Ethnic Background (option to fill in). The racial and ethnic breakdown of the study participants may be seen in Table 3.1. According to HESA (2022), 73% of all UK full-time higher education undergraduate students were white during the academic year 2021-2022. At the university, white students made up 64% of entrants in the same time period. This is just about aligned with the percentage of white students in this study at 68.75%. While 12% of UK undergraduates identify as Asian and 14.16% of the university’s students identify as Asian, 21.88% of this study’s participants were Asian (HESA, 2022). Lastly, 15% of UK undergraduates identify as Black, mixed, or another ethnic minority whereas students in these demographic groups totalled just about 9.37% in this study (HESA, 2022).
Table 3.1: Racial and Ethnic Breakdown of GBV Study Participants

<table>
<thead>
<tr>
<th>Racial and Ethnic Group</th>
<th>Study Participants (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>68.75%</td>
</tr>
<tr>
<td>Asian</td>
<td>21.88%</td>
</tr>
<tr>
<td>Black</td>
<td>3.12%</td>
</tr>
<tr>
<td>Other Racial and Ethnic Background (Middle eastern, not disclosed, etc.)</td>
<td>6.25%</td>
</tr>
</tbody>
</table>

Students were just as diverse in terms of student experience related demographics. 73.68% of the schools were represented, while students came from at least 23 different programmes. A majority 81.25% of students were involved with a university-sponsored activity (ex. sports, society, etc.). Students were asked to report their average number of ‘nights out’, such as going to parties, pubs, clubs, etc. While 31.25% of students reported never going out in their average week at university, 37.50% reported having one night out per week. 12.50% reported having two and three nights out a week each and 3.13% reported having four and five nights out a week each. A similar number of students stated they drink 0 units of alcohol per week on average as those who do not go out, at just about 25.00%. Of the students who drink regularly, 53.13% of students reported 1-7 units of alcohol per week and 6.25% drink 8-14 units of alcohol per week. 15.62% of students reported drinking 15+ units of alcohol per week. Units of alcohol were measured according to the National Health Services’ definition of units of alcohol. This was explained to students with the following examples: one pint of beer is two units of alcohol, while one small glass of wine or shot of alcohol is one unit of alcohol. The UK Government (2021) recommends adults do not drink more than 14 units of alcohol per week, and if they do, to spread it out over at least three days or more. Additionally, according to a survey administered across the UK by the education charity Students Organising for Sustainability (2022), 50% of student respondents reported drinking alcohol at least once a week. This hints at the fact that perhaps the students in this study, and potentially at the partner university, may be more inclined to drinking than the average undergraduate in the UK.

3.2.2: Data Collection

Surveys involved a pre-training survey and a post-training survey. Students completed the post-training survey either within one week post-training or during the third week post-training to assess the longevity of training effects. Survey items used in this study included questions on demographics, awareness and knowledge, bystander expectation and permission, perceptions of community (Unger & Wandersman et al.,
perceptions of safety and trust, the Bystander Efficacy / Confidence Scale (Banyard et al., 2005; modified by Alegría-Flores et al., 2017), the Bystander Intent: Bystander Behaviour Scale- Revised (McMahon et al., 2014), and the Mor Barak Inclusion-Exclusion Scale (MBIE) (Mor Barak & Cherin, 1998). A full list of the survey items are in Appendix B.4. The Bystander Efficacy / Confidence Scale and the Bystander Intent: Bystander Behaviour Scales are both extensively used in the gender-based violence literature, including post-interventions such as in Senn & Forrest (2016) and Amar et al. (2015). While the community scale is not normally administered in gender-based violence studies, this scale was incorporated as it originally was used to measure whether a neighbour is likely to provide emotional support and aid for other neighbours (Unger & Wandersman et al., 1982). Given the sample was from student accommodations and a portion of the focus group was on the student accommodation experience, this scale felt appropriate to measure community (Under & Wandersman et al., 1982). These questions were then adapted to also ask students about their perceptions of safety and trust within their student accommodations and the university at large. Finally, I added the inclusion scale to understand whether there was a relationship between perceptions of inclusion and any of the bystander behaviours.

Focus groups were a key data collection technique from the onset of this study’s methodological design. I deemed this technique appropriate given that student experience, specifically students’ perceptions of trust, is at the heart of the research inquiry. While less common than survey methodologies, other studies such as Htun et al. (2022) have utilised focus groups to assess a gender-based violence intervention as well. Furthermore, focus groups offer more naturalistic settings for participants to reflect and discuss their own experiences (Wilkinson, 1999). No one student’s experience exists separate from other student experiences and gender-based violence is relational after all. While surveys have been critiqued for decontextualising participants from the research inquiry, focus groups create a space where authenticity and co-creation of social context are embraced (Wilkinson, 1999). Focus groups also provide access to group dynamics, further highlighting the opportunity for relational analysis (Seal et al., 1998).

The questions for the semi-structured 1.5 hour focus groups were developed by myself with the help of an undergraduate research assistant to ensure their applicability to the first year student experience. At a high-level, the focus group began with an introduction to myself and the study, a repetition of the at-will participation and confidentiality matters, and a clear acknowledgement of discussion ground rules. Students usually snacked on food during this time as the focus groups were catered, creating an informal and collegial atmosphere. From there, students completed a 10-minute activity to assess their gender preference for flatmates in their student accommodations modelled after a Gaddis and Ghoshal (2020) study. This activity was
an attempt to settle students into starting to think about gender dynamics. A copy of this short activity may be found in Appendix B.5. Students then spent about 20-minutes discussing their student accommodation experiences in terms of safety, trust, and inclusion/exclusion.

The next 45-minutes of the focus groups asked about students’ knowledge, awareness, and behaviours, followed by an original scenario activity and discussion. Scenario vignettes are less commonly used in gender-based violence research, but offer an opportunity to assess how students would behave in a given bystander scenario (Holtzman, 2021; McKnight, 2021). Students were asked to quietly and individually rank the scenarios presented in terms of likelihood of intervening. Scenarios were loosely based off Holtzman (2021) in that they sought to present students with various perpetrator identities, victim identities, bystander identities, alcohol presence, and awareness of intent. For a copy of these four scenarios, please see Chapter 4.1. It is worth noting that while trust was only directly asked about once during the focus group, it was indirectly a topic of a few other questions and overall a recurring theme. All focus groups were recorded and initially transcribed using Otter.ai.

As mentioned, the 60-minute individual 1:1 interviews were not originally part of the research design. While previous studies such as Acquaviva et al. (2022) and Stojanov et al. (2021) have used individual interviews to measure gender-based violence intervention effectiveness, we did not think we would have the capacity to institute them for what we expected to be a large sample size. On the contrary, once we decided that they were feasible for the smaller Cohort 1, it was clear that they provided an opportunity for the research team to explore some emerging themes from the focus groups more in-depth (Seal et al., 1998). A full overview of participants numbers per research stage may be seen in Table 3.2. Conducted virtually as it was summer holiday for students, the participants were asked questions pertaining to the training intervention reflections, their ideal gender-based violence strategy, ‘the line’ that defines decisions to intervene, and their past bystander experiences. Each of these topics were emerging themes from the focus groups and they were decided on upon consensus from the PhD researcher and research assistant. All individual 1:1 interviews were recorded and initially transcribed using Otter.ai.

The response to the in-person focus groups and virtual 1:1 interviews was overwhelmingly positive in terms of engagement and direct feedback. Likely, this is since the researchers had ample time to prepare for facilitation. I reviewed interview protocol extensively with the two research assistants ahead of their participation in co-facilitating any focus group or interview portions. A script and protocol were drafted weeks ahead of time. Furthermore, I have experience with facilitating focus groups and interviews in my previous work as a human resources professional and as a diversity and inclusion consultant. The informal rapport with students was well established and
likely helped by the fact that a similarly-aged research assistant was usually present. Focus groups are considered helpful in shifting the power from researcher to participant. As a result, this allows participants to feel more in control of the direction of the conversation (Wilkinson, 1999). This power shift was also likely due to the similar ages and life stages, and the fact that the overall atmosphere was informal with food and beverage being consumed throughout. One student, upon being asked who the ideal facilitator for an intervention would be reflected, “[…] if it’s just a random student […] it would be a bit weird, because there’s not connection to like a kind of institution […] if it’s just a member of staff that it feels too authoritative. Because the good thing about the focus group was you guys are like students, so like, you’re like, friendly.” Additionally, students did not seem bothered by the recording devices in the focus groups. These were highly visible and consisted of either an iPhone, iPad, or laptop. Two were usually used at once to ensure dialogue was accurately captured. The recording device in the virtual 1:1 interview was non-visible, and pre-established rapport with the participants was clear, only adding to the authenticity of these conversations.

Table 3.2: Cohort Participant Numbers Per Research Stage

<table>
<thead>
<tr>
<th>Research Stage</th>
<th>Cohort 1 Participants</th>
<th>Cohort 2 Participants</th>
<th>Total # of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expressed interest in the study via email or QR code signup</td>
<td>28</td>
<td>56</td>
<td>84</td>
</tr>
<tr>
<td>Sent the consent form after confirming eligibility for study</td>
<td>27</td>
<td>47</td>
<td>74</td>
</tr>
<tr>
<td>Signed the consent form</td>
<td>11</td>
<td>23</td>
<td>34</td>
</tr>
<tr>
<td>Completed pre-training survey #1</td>
<td>11</td>
<td>21</td>
<td>32</td>
</tr>
<tr>
<td>Completed Epigeum modules (Consent Matters for Cohort 1, Consent Matters &amp; Tacking Harassment for Cohort 2)</td>
<td>9</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>Completed in-person focus group (or online survey equivalent)</td>
<td>9</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>Completed post-training survey #2 (either one week or three weeks post-training)</td>
<td>8</td>
<td>10</td>
<td>18</td>
</tr>
<tr>
<td>Completed individual 1:1 interviews (6-weeks post-training)</td>
<td>7</td>
<td>N/A</td>
<td>7</td>
</tr>
</tbody>
</table>
3.2.3: Limitations

The study had a small sample size despite an initial expectation, based on previous student appetite, that a larger sample would be a given. From an empiricist perspective, the study’s small sample size would call into question the study’s ability to draw generalisable inferences. However, as shown above, the sample is fairly representative of the average university student. As discussed above, participants’ gender, race/ethnicity, and drinking habits breakdown were very on par with statistics from either the university itself or other research. Furthermore, while the small sample size was unexpected, a benefit of it was that it allowed us time to add 1:1 interviews for Cohort 1. This in turn allowed for a more intimate exploration of the student experience than what otherwise would have been possible if we had recruited the original number of 240+ students we intended to recruit (Crouch & McKenzie, 2006).

I find that the small sample size, despite the university’s best effort to recruit, is a finding in and of itself. I believe that the difficulties in study recruitment were representative of a larger issue at the university. With that, despite a hefty bureaucratic battle to get an increased incentive approved, it seems that the increased incentives for Cohort 2 did not have a major effect on the number of participants who completed the study. Likely, the incentives in combination with an increased amount of online and in-person signposting, did lead to exactly double the number of students expressing interest in the Cohort 2 study. Even so, there were students who did not qualify for Cohort 2 (seven second-year students, one commuter student who did not live in university student accommodation, and one postgraduate).

While in Cohort 1, timing of the study was an obstacle for students, it seems this was less of a hurdle for students in Cohort 2. Only two students cited timing or university workload as reasons they did not want to participate. It is worth noting that the largest drop off throughout the study was seen during the training administration stage. A total number of 10 students did not complete the training and thus dropped out of the study. Only one student wrote in to say that they could not complete the training due to increased workload during the academic term. While Cohort 1 students came together prior to the focus group to complete the training in the presence of myself, Cohort 2 students had just over a weeks’ time to complete the training unsupervised. This brought the total number of Cohort 2 students who were eligible for continuing in the study to a mere 11 students, just two more students than at the equivalent stage for Cohort 1. Related to this, participants who did continue with the study were likely to be students who were either especially free, or, potentially passionate about the subject matter.

Altogether, I, along with the research committee, fully expected the increased incentives to significantly boost study interest and participation numbers. Tables of
final participant numbers per survey and focus group may be seen in Appendix B.6. As seen from this discussion, this was not the case. It seems likely that asking students to complete the 3.5-5 hours total training modules in their own time during the term is ineffective, even with the increased incentives. This limitation of a small sample size is therefore a major finding for the university, as it is evident that students cannot be depended on to complete such a long training without it being mandated. This is even the case when it is a training that originated from an ask from the largest and most representative student association on campus. In my final internal report given to the university, I stressed how the difficulty of recruitment, even with incentives, was enough of a reason to mandate the training in the future. For more information about the setbacks to and incentives for this study, please see Appendix B.3.

3.3: Inclusion Study Research Design

The sampling method for the inclusion study was also purposive. The two participant schools, School A and School B, were the two schools at the university where inclusion training was made mandatory for some tutors and introductory tutorials were offered to first year students. There was one other school where inclusion training was mandatory, but the school did not offer an undergraduate programme at the time. Each tutorial had about 10-20 students each, and the tutorials were 50-minutes in duration led by one tutor. As the classroom is made up of not just students, but also the one tutor, these were together the study’s research participants. Participants partook in two inclusion surveys, as well as three recordings and/or observations, dependent on if their full tutorial group consented to being recorded and/or observed.

The research design was comparative in nature, but this was not the main focus of the research. While the comparison stemmed from the fact that the type of training tutors underwent was dependent on which school they were in, the main reason for studying across the two schools was that it allowed insight into university-wide inequality regimes. Studying the two schools helped show how inequality regimes manifest across the university’s different demographic groups and schools. Given the university’s diversity and inclusion strategy is highly school dependent, it is worthwhile to see how inequality regimes may persist regardless of the school-specific strategy. At the same time, the research design did allow for some comparisons between the schools to be drawn given each school had a different inclusion training. The main differences between the trainings were that School A’s was focused on case studies, while School B’s was more lecture-style. School B consequently also focused more on the foundational knowledge of what is diversity, equity, and inclusion while providing tutors with tips on creating inclusive classrooms. School A instead allowed more space
for tutors to contribute their own knowledge and share how they would approach different situations.

More specifically, School A’s training was heavily case study based whereby groups of tutors were given common inclusion/exclusion scenarios and discussed how to address the situations. For instance, if a student is deaf and is put into a group discussion, but other students struggle to understand deaf voice, how should a tutor inclusively handle that? Another scenario addressed what steps a tutor should take if a student claims a tutor is being discriminatory in their marking practices. School A’s training was developed by one person who is a professor and the director of equality, diversity, and inclusion for that school. The scenarios came from their personal experience as a lecturer. It was 2-hour long and delivered both in-person and virtually by the professor. The training was not new, as it had previously been facilitated the academic year prior. It was not mandated for all the school’s tutors, but it was mandated for the tutors on the introductory course. Tutors who attended were compensated for their time.

School B’s training was developed by a working group of tutors, including me. We worked closely with the relevant school’s teaching development supervisor to ensure the training had subject matter expertise and relevant institutional knowledge. This was the first training of its kind to be mandated for all tutors in the school. Two in-person 2-hour training sessions were facilitated by me and another tutor from the working group to 120+ tutors in the school. New tutors and returning tutors for the 2022-2023 academic year were all mandated and compensated at the standard hourly rate to complete the training. The training focused on teaching tutors how to alter behaviours to promote classroom inclusion, as well as increasing awareness and knowledge of equality, diversity and inclusion. With that, about 20-minutes were spent ensuring tutors knew how the university defines diversity, equality, and inclusion. The remainder of the time alternated between small group discussions and lecture-style review of best practices when it comes to inclusive classrooms. We taught tutors what they can do before the course begins, such as reviewing the syllabus for missing perspectives. We then discussed what tutors can do throughout the course, such as critiquing the canon, checking in with students who may not be participating, incorporating accessible design, and alternating classroom seating arrangements.

The research design also encountered several setbacks and challenges. While the initial proposal was submitted to the ethics board in July 2022, and received approval in August 2022, the design had to change significantly by September 2022. The original design failed to realise how frequently students, especially first year students, were still going to be encountering schedule changes well into the term. Schedule changes meant students were being added and removed from the introductory courses’ rosters daily. Of course, this made obtaining consent from these
students rather difficult. To record tutorials, I needed full consensus from every student. This made it extremely difficult to gain full consent from a tutorial group for recording unless the tutorial group remained the same. Furthermore, this issue was exacerbated by a major glitch in the university’s online scheduling software. Students were entirely deleted from one course’s student roster, then re-added to a different tutorial group even if they did not request a schedule change. For several weeks, tutors were unaware of which students were actually in their classes, while students’ names sometimes were on the student list despite the student having dropped the course entirely. All of this meant that the start of data collection, and resulting data collection points were pushed back further into the term than originally planned. There were still a few names on the official student roster by end of term who were not actually enrolled in the course.

Consent for this research was split into three consent “tiers”, as outlined in Table 3.3. Tier I consent was when participants agreed to fully participate in the research, including both filling out two inclusion surveys, and having their tutorial both recorded and observed at three different points in the term. Given some participants were uneasy with being observed in their classrooms, Tier II consent meant participants filled out the surveys and agreed to being recorded, but not observed. Of course, not all participants in a given tutorial group agreed to being either recorded nor observed in the classroom. In this case, participants fell into a Tier III consent level where they just filled out the inclusion surveys. While neither tutors nor students were monetarily incentivised to participate in this research, they were informed that this research was critical to allowing the university to develop its equality, diversity, and inclusion strategy. Furthermore, tutors were promised high-level feedback on how they could better foster inclusion in their tutoring. For a review of how this study was communicated and a copy of consent forms, please see Appendix C.1 and Appendix C.2.
### Table 3.3: Inclusion Study Research Design for Consent Tiers

<table>
<thead>
<tr>
<th>Timeline</th>
<th>Research Stage</th>
<th>Tier I</th>
<th>Tier II</th>
<th>Tier III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Data Collection</td>
<td>Express interest in the study</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Sign the consent form after eligibility for study confirmed</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Start of Course Term</td>
<td>Complete inclusion survey #1</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Participate in tutorial recording and/or observation</td>
<td>X</td>
<td>X</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td><em>(Observation &amp; Recording)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>(Recording only)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mid-Point of Course Term</td>
<td>Participate in tutorial recording and/or observation</td>
<td>X</td>
<td>X</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td><em>(Observation &amp; Recording)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>(Recording only)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>End of Course Term</td>
<td>Complete inclusion survey #2</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Participate in tutorial recording and/or observation</td>
<td>X</td>
<td>X</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td><em>(Observation &amp; Recording)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>(Recording only)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.3.1: Participants

By the end of the recruitment process, three tutors from School A consented and four tutors from School B consented. All tutors consented to being both recorded and observed, except for one who knew me on a more personal level. It was decided the tutorial groups of that tutor should not be observed, but just recorded. This resulted in students being recruited from 10 tutorial groups in School A and 17 tutorial groups in School B. At the official start of data collection, 256 participants were sent the first survey. Of the students who consented, just seven agreed to only being recorded and not observed. The number of student participants fluctuated over the following weeks as students switched timetables, dropped out of the course, and joined the course. The second and final survey was sent to 255 participants showing that despite fluctuations in tutorial enrolment, the total number of student participants stayed about the same. No tutors dropped out of the study. Unfortunately, it was difficult to get tutors to fully consent to being recorded and observed. Even if one student did not consent, that meant the tutorial could not be recorded and/or observed. In the end, only one tutorial from School A was observed and recorded, while two tutorials from School B were recorded only.

Demographics of the student participants will be discussed, but the demographics of the tutors will not be shared at a high-level given the small sample size (n=7). 148 students filled out the first survey and 115 filled out the second survey, of which 10 never completed demographic questions. Those students were disregarded in analysis except for when taking full averages. The reason they did not complete demographic questions is because they only completed the second survey which did not have the demographic questions again. This means 166 unique student participants completed a survey. Students were asked about their age, school, programme, gender, sexuality, disability status, and involvement in university activities.

The average participant in the study was 18.70 years old and identified as a heterosexual woman with no disabilities. 50.64% of participants were 18 at the start of data collection, but ages ranged from 16 to 41. 76.92% of participants were women, 19.87% were men, and 3.21% were non-binary and/or transgender. This is substantially different than the overall percentage of women across the university which is about 61.55% female and 38.18% male. At the same time, it is well known that at least one of the courses, the one in School B, is female dominated (67.7% for entrants in 2021/2022). 75% of students identified as heterosexual, while 20.51% identified as non-heterosexual. 4.49% of student participants identified chose not to disclose their sexuality. 91.67% identified as not having a disability and 8.33% identified as having a disability, which is lower than the university’s reported 17.10% of undergraduates with a disability. It was about an even split between students who were involved with university
sponsored activities (51.28%) and those who were not (48.72%). This is to be expected as 81.41% of students were in their first year of university. Interestingly, there was also a relatively even split between students who were officially in the participant schools (52.56%) and those who were enrolled in a different school (47.44%). This is interesting because as schools try to target higher levels of inclusion for their students, training interventions like tutor training may not be the most effective route to reach their own student cohorts. In fact, 59 different degree programmes were represented.

As seen in Table 3.4, of the students who submitted demographics, 56.02% were white. This is much lower than the average percentage of white undergraduate students in full-time higher education which was 73% in the academic year 2021-2022 (HESA, 2022). The university reported 64% of students as white identifying for the same time period. Just as with the gender-based violence study, Asian students were overrepresented in this study at 31.93%. This is compared to 14.16% of entrants at the university being Asian identifying and 12% of all UK undergraduates identifying as Asian (HESA, 2022). Of the Asian students in the sample, 71.15% were Chinese students. This is aligned with the university’s own demographics as Chinese students are the largest international student population at the university.

Table 3.4: Racial and Ethnic Breakdown of Inclusion Study Participants

<table>
<thead>
<tr>
<th>Racial and Ethnic Group</th>
<th>Study Participants (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>56.02%</td>
</tr>
<tr>
<td>Asian</td>
<td>31.93%</td>
</tr>
<tr>
<td>Middle Eastern</td>
<td>3.01%</td>
</tr>
<tr>
<td>Other Racial and Ethnic Background (ex. Black or Latinx)</td>
<td>3.01%</td>
</tr>
</tbody>
</table>

3.3.2: Data Collection

I chose surveys as a form of data collection as they are a consistent way to quickly understand all participants’ perceptions of inclusion at different points in the term. It is important to note that I did not administer a pre-training survey given consent could not be gathered in time. This means the research design cannot directly assess the training intervention’s influence, but the surveys can investigate changes over time of repeated measures. I selected the Mor Barak Inclusion-Exclusion Scale as it has been routinely validated to measure perceptions of contribution and participation in work group processes (Mor Barak and Cherin, 1998). One school’s course organiser was adamant that surveys should not be longer than 15 minutes including demographic questions. While it would have been ideal to also add questions to ascertain
participants’ perceptions of uniqueness and belongingness in their tutorials, there was simply not enough time. Instead, I felt it was most important to collect students’ demographics and perceptions of contribution and participation to compare these to patterns of contribution and participation behaviours in the recordings.

The Mor Barak Inclusion-Exclusion Scale questions were adapted to the classroom setting as the scale was originally developed for workplace settings. Furthermore, perceptions of community (Unger & Wandersman et al., 1982), and perceptions of safety and trust were also measured for, just as in the gender-based violence study. These measures were an attempt to further understand the student experience within the classroom and how it related to their greater university student experience. Demographic questions were also asked. These were all combined into one survey, found in Appendix C.3, which was given at the beginning of the tutorial term and the end of the tutorial term. I calculated the Cronbach’s alphas which demonstrated good reliability at 0.81 and 0.80 for the adapted scale for the beginning and end of term surveys, respectively. Previously, studies utilising the standard MBIE Scale reported Cronbach’s alphas of 0.88, 0.90, 0.81, 0.87, 0.81, 0.84, and 0.84 (Brimhall et al., 2017; Mor Barak, 2013).

The data collection points allowed me to understand levels of inclusion, safety, community, and trust at the beginning and end of the tutorial lifecycle. The surveys were administered during tutorial #4 and #9 for both courses. This corresponded to academic Week #5 and Week #10 for School B, and Week #5 and Week #11 for School A. This is because School A had a one-week break mid-term, while School B did not. I sent the surveys the same day as the tutorial, and they remained open for exactly one week. Participants could fill it out up until the next tutorial. I sent participants at least two additional email reminders to submit the surveys.

Tutorials where participants fully consented were recorded and/or observed during three touchpoints in the term. Recordings were initially transcribed using Otter.ai. As the survey offered insight into perceptions of inclusion/exclusion, recordings and observations allowed me to monitor behaviours of inclusion/exclusion. For School A, this corresponded to tutorials #4, #6, and #9 in Week #5, Week #8, and Week #11. For School B, this corresponded to tutorials #4, #6, and #9 in Week #5, Week #7, and Week #10. For the tutorials that were only recorded, an iPad was used. The location of the iPad in the classroom depended on the layout of the classroom and where students were sitting on that given day. Usually, this was in the back of the class. For the observations, I sat in the back of the classroom with a mask on while taking notes. The iPad was always next to me to ensure I was not included in the recording. These recordings and observations were chosen as data collection methodologies because of their ability to generate data that could be used for social network analysis. The observations were critical in helping me understand the group dynamics of the tutorials.
firsthand to later guide analysis. For instance, while it was known that social network analysis would be used, I was able to generate ideas for this analysis while observing students. After all, the ebbs and flows of communication to and from marginalised groups is crucial in understanding how inclusion shifts to redistribute power (Helgesen, 1995).

At first, I personally did not consider these observations and recordings a form of ethnography. However, once data collection began, it was clear that I was partaking in a reflexive ethnography rather than purely a field observation. I was not just listening in on the tutorial content and taking the sessions at face-value as one would in a field observation. Rather, I was intensively engaging as an ethnographer with the tutorials, reflecting on the setting, the class seating charts, the dialogue, and the rituals (Ciuk et al., 2018). I was actively taking into consideration the larger social context of the university and the country in which the university is located (Ciuk et al., 2018). Accordingly, the recordings acted as a virtual ethnography or a study of micro-interactional dynamics (Fuhse, 2023). I could dissect the importance of the mundane, such as a pat on the back, an interruption, and other body gestures that play into student-to-student and student-to-teacher dynamics in the classroom (Ciuk et al., 2018; Fuhse, 2023). While I was able to observe the class time either face-to-face or virtually, I was unable to observe the digital interactions between students and tutors. Akemu & Abdelnour (2020) stress the importance of digital interactions such as emails, Whatsapp groups, and social media encounters. The data collection did not include such digital artefacts. I thus do not claim to fully understand the participant tutorial experience. Even so, the ethnography helped point me in the direction of what types of analyses make the most sense to begin to understand the participants’ experiences of inclusion/exclusion within this glimpse of their overall university experience.

3.3.3: Limitations

A few limitations affected this study. While there is a limitation in terms of the small number of tutorials that fully consented to either being recorded or observed, I do find that my overall sample is relatively representative of the university’s demographics in terms of disability and race/ethnicity, even with a slight over-indexing of Asian students. Similarly, the sample can be considered representative in terms of gender given the female-dominated nature of at least one of the schools. Additionally, while I had hoped more tutorials fully consented to either being recorded or observed, I found that I’m unsure if I could have handled any more fully consenting. My time during autumn 2022 was completely limited by simultaneously running both the studies. I think the larger limitation, which may be a finding in itself, is that although the tutorial sizes were between 11-14 students on the official student attendance roster for each tutorial,
actual student attendance was much lower. This is critical for the university to know, especially as the university tries to navigate a “post-COVID” push for in-person courses. Low attendance was common nearer to the end of term and can be seen in the attendance for the one tutorial I observed. Given student examination schedules, student attendance dropped to a lowly five by the last tutorial date for the observed tutorial in School A. While this means the analysis of student interactions that can be done on observations is limited, it does also present an opportunity to see how perceptions of inclusion/exclusion are impacted, if at all, by participation. Across all tutorials, strikes and exam schedules influenced student attendance. One other major reason for attendance drop-off emerged too. In one informal conversation after the final tutorial date, a tutor reflected on the fact that it was clear the FIFA World Cup may have played a role in student attendance. The competition ran from November 20th - December 18th, 2022, which perfectly aligned with the end of term. As stated, student absences in tutorials may be seen as a learning in and of itself, as this study does conceptualise participation as a form of inclusion/exclusion.

3.4: Analysis

Schwemmer & Wieczorek (2020) found that historically, sociology maintains a distinct methodological divide whereby researchers are either quantitative or qualitative in their chosen methods. Ironically, through their purely computational and quantitative study, the researchers found that sociology is becoming, slowly, more quantitatively driven (Schwemmer & Wieczorek, 2020). For this reason, this mixed-methods research study is a minority in the field of sociology in that it will not define itself by either/or, but rather by both quantitative and qualitative analysis. The ask to the reader of this thesis is to embrace methodological eclecticism. Coined by Tashakkori & Teddlie (2010), methodological eclecticism is when a researcher chooses the methods of analysis that best answer the research questions, therefore deconstructing the methodological divide that plagues sociology.

In following methodological eclecticism, I hope to join Dimaggio (2015) and Abbott (1998) in helping social scientists relax their obsessions with the gold standard of causality. This research and its analysis are exploratory in nature. It is my attempt at bridging the qualitative and quantitative methodological divide. Simply put, I like numbers and I like people. Like Tashakkori & Teddlie (2010), I do not wish to put people and numbers in a fight against each other. With that, I have chosen to use thematic analysis, statistical survey analysis, computational text analysis, and social network analysis. Each of my four findings chapters will adopt one of these methods as its main methodology. I provide a high-level overview of each below. Given this chapter already must balance two studies and review four methodologies, each findings chapter will
also include a more detailed exploration of my methodological decisions more in-depth. I end this section engaging in some methodological reflexivity given no analysis is neutral (Corlett & Mavin, 2019).

3.4.1: Thematic Analysis

Thematic analysis is a rather straightforward, but flexible qualitative method used to identify patterns within a given dataset. In this case, thematic analysis will be used on focus group conversations whereby the dialogue has been converted into text data. The researcher plays an active role in thematic analysis and chooses which patterns are worthy of further analysis; it is a process, like many methodologies, filled with bias and human judgement (Braun & Clarke, 2008). The research utilises thematic analysis that is inductive and latent in nature, in that I tried my best to not let theory guide the coding. I also explored hidden meanings of the data. This is because thematic analysis is used in this research to explore trust in gender-based violence through a pluralist framework. Other related studies such as Holtzman (2021) have used a version of thematic analysis in their own analysis. However, thematic analysis was not previously mapped out step-by-step in a replicable, but flexible way until 2008 when Braun & Clarke (p. 87) published the following key phases of thematic analysis:

1. Familiarise yourself with the data: Transcribe data (if necessary), reading and re-reading the data, noting down initial ideas.
2. Generating initial codes: Coding interesting features of the data in a systematic fashion across the entire data set, collating data relevant to each code.
3. Searching for themes: Collating codes into potential themes, gathering all data relevant to each potential theme.
4. Reviewing themes: Checking if the themes work in relation to the coded extracts and the entire data set, generating a thematic ‘map’ of the analysis.
5. Defining and naming themes: Ongoing analysis to refine the specifics of each theme, and the overall story the analysis tells, generating clear definitions and names for each theme.
6. Producing the report: The final opportunity for analysis. Selection of vivid, compelling extract examples, final analysis of selected extracts, relating back of the analysis to the research question and literature, producing a scholarly report of the analysis.
While I am claiming the thematic analysis process I am undertaking is inductive, rather than deductive, this of course comes with a major callout. At the time of analysis, I had spent the last 2.5 years reading gender-based violence literature. It is inevitable that gender-based violence theory directly impacted my thematic analysis, from the initial jotting down of my ideas to the chapter production. Theory of course played a major role in my thematic analysis, but in claiming it was inductive, I did not let any one theory drive my coding. Where I find that certain theories are clearly emerging in the development of my codes, I explicitly state their influence in the analysis. It is also worth stating that I intentionally completed my thematic analysis prior to computational text analysis of the same data. This was an attempt to eliminate any bias carrying over from my computational text analysis findings.

### 3.4.2: Statistical Analysis

Where appropriate, my research also incorporates statistical analysis of surveys and namely of Likert scale data. A Likert scale is an ordinal scale used to measure an individual’s perceptions of attitudes, opinions, and behaviours. Usually, this is conducted by presenting a written statement to an individual and asking them to what extent they agree or feel strongly about the statement. In this case, I incorporated two Likert scales into both the gender-based violence study and the inclusion study. Depending on the scale used, individuals were either asked to rank their opinions from 1-5 or 1-6.

Statistical analysis of Likert scale data is a highly debated topic. Some statisticians feel that if a dataset has a small sample size and non-normal distribution, parametric statistics cannot be used for analysis (Norman, 2010). Parametric tests generally require continuous interval data, normal distributions of that data around the mean, and only 5% of results can fall beyond two standard deviations of the mean (Mircioiu & Atkinson, 2017). Non-parametric statistical tests do not maintain any assumptions about the results’ distribution around the mean and can be used with ordinal data such as Likert scale data (Mircioiu & Atkinson, 2017). Other passionate statisticians argue for and have proven that with a sample size of at least 15, most parametric and non-parametric tests will conclude the same results for non-normal data (Norman, 2010).

As a result, I employ different statistical tests for my data depending on the study. For the gender-based violence study and its non-normal small sample, I decide to rely mostly on non-inferential descriptive analysis. Where applicable, I do bring in some non-parametric analysis, but I do not draw major conclusions from these statistical tests. For the inclusion study, I rely on both non-parametric and parametric statistical analysis. For individual questions on the Mor Barak Inclusion-Exclusion scale, I use
Mann-Whitney Wilcoxon and Kruskal-Wallis non-parametric tests. To conduct parametric tests, I sum the MBIE scores into a cumulative MBIE score to produce pseudo-continuous interval data. I then utilise the Welch two-sample t-test and Two-Way ANOVA, depending on if the subgroup has more than two categories (ex. sex, race and ethnicity, year in school). I also occasionally use a Chi-squared test and outline explicitly when I do.

3.4.3: Computational Text Analysis

Analysing text in sociology is not uncommon, but doing so quantitatively is (Schwemmer & Wieczorek, 2020). For example, the text data from focus group data is generally analysed for content, rather than dissecting the interactions themselves (Wilkinson, 1999). These analyses are then generally paired with survey data, which this study will be doing too (Seal et al., 1998). As Goldberg & Srivastava (2022) point out, there is an opportunity to assess a field site’s climate as a whole through text data, rather than solely administering routine surveys. This study will do just that by measuring trust in the dialogue of participants’ focus groups and individual interviews within the gender-based violence study. My decision to work with computational text analysis, after completing thematic analysis, is to show how computational methodologies may provide a more reproducible alternative to thematic analysis. At the same time, I chose computational text analysis in the hopes that it may confirm and add to my conclusions reached through thematic analysis. Therefore, I hope to show how computational text analysis can be seen to be both a replacement for and complementary to thematic analysis. In doing so, I explore whether a blueprint for incorporating computational methodologies into gender-based violence literature exists.

The general process for computational text analysis that will be followed is that of Benoit (2020). Just as with thematic analysis, this process is highly iterative, embodying the exploratory nature of this research once again. The data to be analysed are unstructured, as text is pulled from the recordings of dialogue (Benoit, 2020). It is critical to underscore that while the text data are from dialogue, of course the participants did not intend for it to be analysed so exhaustively as the process below entails. This is the nature of computational text analysis though, where data are broken up until it is inscrutable, then meaning is drawn from it again through an entirely different lens. Benoit’s (2020, p. 468) process for computational text analysis is thus as follows:

1. Selecting texts and defining the corpus.
2. Converting the texts into a common electronic format.
3. Defining documents and choosing the unit of analysis.
4. Defining and refining features.
5. Converting textual features into a quantitative matrix.
6. Analysing the data using an appropriate statistical procedure.
7. Interpreting and reporting the results.

In this case, my statistical procedures will be using both word embeddings and sentiment analysis. My decision to use word embeddings is guided by the fact that I hope to develop a bespoke trust dictionary to run sentiment analysis. Word embeddings are vector representations of words that frequently co-occur within text data (Li et al., 2020). At a high-level, I hope to first explore what trust words are used in gender-based violence research and what words they occur with. I then will apply that list of words to my study’s data to quantify trust levels within the student population. This will be done through sentiment analysis. Sentiment analysis involves comparing words from text data to a setlist of words that make up a dictionary (Pennebaker et al., 2007). The frequency of matches between the text data words and the dictionary words shows to what extent either a tone or emotion is present (Grimmer & Stewart, 2013). A limitation of sentiment analysis using dictionaries is that when applied to a context for which it was not developed, the results can be grossly misrepresentative (Grimmer & Stewart, 2013). For this reason, I am choosing to develop my own dictionary using word embeddings to overcome this lack of validity. While Pennebaker et al. (2007) used emotion scales, thesaurus, and a Webster dictionary, to develop the LIWC dictionary, I am solely focusing on gender-based violence questionnaires and word embeddings to develop my trust dictionary. I will largely build on Alsaid et al.’s (2022) work and compare my bespoke dictionary to two other pre-existing dictionaries. More details about my computational text analysis decision-making such as pre-processing, choice in dictionary, and analysis steps will be included in Chapter 5.

To further boost the validity of my computational text analysis, I will cross-compare my results with thematic analysis. Grimmer & Stewart (2013) discuss how a limitation of computational text analysis is that it does not reveal the nuance and context of what close reading uncovers. In the case of my study, computational text analysis will not reveal the reasons for certain levels of trust and/or distrust. Rather, this comes with close reading through methods such as thematic analysis. It is for this reason why I will also use computational text analysis in addition to thematic analysis. The pair go hand in hand together to boost my analysis of trust as it relates to gender-based violence within academia. It is my hope that this study will showcase how computational methods may help paint a full picture in a way that one method alone cannot.
3.4.4: Social Network Analysis

Social network analysis (SNA) is the empirical study of networks, with the focal point of study being the relations between individuals or groups of actors (Crossley, 2016). SNA makes the most sense to use in this research study as the method scrutinises not just any one individual’s experience, but the communal experience at large in a given field site. In this way, the method is a natural and computational extension of relational sociology in that it can reveal hidden power structures and differences in capital between actors in a network (Crossley 2016; Wagner & González-Howard, 2018). The network in this research study is the classroom tutorial group, where each actor represents one individual; either a student or a tutor. Wagner & González-Howard (2018) have loosely outlined how to apply SNA to a classroom setting, with the steps below:

1. Choosing a network to study.
2. Defining who or what constitutes an actor.
3. Defining what type of interaction counts as a relational tie.
4. Creating a matrix of actors and ties.
5. Analysing the network as a whole.

Using SNA to study classrooms is not new, but it is highly underutilised in the studying of inclusion/exclusion (Wagner & González-Howard, 2018). While Karimi & Matous (2018) used it to study student inclusion in student societies at a university, and Mameli et al. (2015) used SNA to measure primary classroom interactions, this study will be the first of its kind to use SNA to measure in-class inclusion at the university-level with demographic analysis and survey analysis too. Just as with thematic analysis and computational text analysis, the process of implementing social network analysis will be iterative and exploratory. While defining an actor is simple, the analysis will be a time-consuming exploration of which ties prove significant to study. Even so, I am choosing to pursue social network analysis to measure inclusion to show whether computational methodologies provide a more reproducible and complementary alternative to more routine survey analysis. As computational text analysis will be used to ideally confirm and add to thematic analysis, social network analysis will be used to build upon and hopefully confirm the results of survey analysis. In this way, it is my hope that the more computational method can be utilised to confirm and add nuance to the findings of the less computational methods. In particular, with social network analysis being used to measure inclusion, I can see how behavioural measurements of inclusion, defined as participation and contribution, can be compared to participants’ perceptions of participation and contribution. This offers a blueprint for how inclusion
researchers can move the inclusion literature forward by adding computational methodologies into their toolkit.

To carry out social network analysis, I will record certain tutorial groups whose participants fully consented to participating in the study. I will carry out three recordings per tutorial group over the term. I will convert these recordings into an adjacency matrix to quantify interactional ties, determined as one individual speaking to another individual or the full class. From there, I will analyse both individual class networks and aggregate tutorial group networks over the full term. This will both provide a snapshot into particular class sessions, as well as help me understand how inclusion shifts, if at all, over the 10-week term. Methodologically, the end goal is to see how various methods can build upon each other. The empirical goal is to understand how inequality regimes may subtly exist in terms of inclusion/exclusion and if there is more that can be done to improve inclusion levels within tutorial groups. I will provide further detail in Chapter 7, especially as it concerns which social network analysis metrics I will study, and how they may look in a hypothetical and perfectly inclusive/exclusive network.

3.4.5: Limitations of Methodologies

While the limitations of thematic analysis and statistical analysis of surveys are relatively well known in the sociology community, I would like to discuss a couple of them here. Primarily, I have already touched on the role of theory in my inductive thematic analysis process. To add to this, I want to acknowledge how my thematic analysis will of course be based on my judgement and inevitably influenced by my three years of reading gender-based violence literature. In this way, the language I use is limited in some ways. Although I may write of themes ‘emerging’ or ‘uncovering’ conclusions, I want to explicitly share how I recognise that this language falsely identifies my conclusions as the conclusions (Braun & Clarke, 2006).

With that, even though I use this language in the subsequent chapters both in my thematic analysis and beyond, I acknowledge that my findings are surely not the only findings and will be limited in their widespread applicability for this reason. In much the same way, I would like to quickly discuss the limitation of statistical analysis of surveys. While surveys are helpful in drawing out perceptions from large groups at one moment in time, there may still be some uncontrolled variance. I did not define in the surveys key terms like community, trust, safety, participation, contribution, and inclusion for participants. Therefore, participants may have answered questions differently based on how they interpret some of these terms and their applicability to their university experiences. If this study were to be reproduced, it may be worthwhile to first understand how participants define these terms, or incorporate a definition for participants into the surveys themselves.
It also feels important to emphasise some limitations to the computational social science methodologies I chose. As already stated, the vast majority of student interactions will not be captured as tutorials are just one facet of student life. The rest of their student experience, namely, societies, extracurriculars, student accommodations, and ‘off-campus’ personal lives will not be part of this study. This illuminates the pressure to collect ‘enough’ big data, without invading participants’ personal privacy (Collins & Steffen-Fluhr, 2019). Other non-tutorial interactions between students will also not be observed (for example, if students are texting each other during the tutorial, or chatting before or after the tutorial officially begins). From a researchers’ perspective, it would of course be endlessly rich to monitor participants’ day-to-day interactions. However, I find I was able to strike an appropriate balance of monitoring and privacy by monitoring just three tutorial sessions. This allowed for students to settle into their tutorial groups without constant observation. Similarly, the focus groups in the gender-based violence study are an artificial group where students may be able to reflect on their trust in other students, but trust they experience in friend groups and groups of student strangers may fluctuate over the academic year. My hope in conducting the focus groups was to discuss trust, without explicitly asking about solely trust. This allowed for trust, and distrust, to become a more naturally occurring theme throughout the discussions if students felt it was relevant. Additionally, given exclusion and distrust are present in arguably every group interaction in some form, I can only hope that the combination of surveys and computational social science methodologies can better guide the university towards understanding how to reach higher thresholds of inclusion and trust, while mitigating exclusion and gender-based violence in the student experience.

3.5: Ethical Considerations

In using computational data to study gender-based violence and inclusion/exclusion, this study could lead to recurring violence of ‘othering’ historically marginalised groups, as well as putting these categorical groups at even more risk of trauma (Hoffman, 2019). A major ethical consideration related to this is how to anonymise data and protect individuals, especially those from historically marginalised backgrounds. Video and audio recordings will be kept on a password-protected cloud, only accessible by the PhD researcher. They will be destroyed once transcribed and fully analysed. Anonymisation and protection does not end there. One question I, and other researchers, have grappled with is how to categorise individuals in terms of their dimensions of diversity (Gebru, 2020). I do not want participants to feel tokenised or at risk of identification, especially as tokenisation is a form of othering and is “as disempowering as complete exclusion” (Crenshaw, 1991). Yet, I also do not want to
categorise individuals into larger, more anonymising demographic groups, and consequently ignore participants’ identities. This has taken careful consideration with each and every analysis run, especially when it comes to age, gender, sexual orientation, race, and ethnicity. These dimensions of diversity often have minority groups that are small, allowing for easy identification. The way I have treated these categorical groups in the following chapters, I hope, shows participants that I wish to protect them all the while not completely erasing their narratives and only adding to the potentially already traumatic experience of navigating historically oppressive spaces.

Ahead of beginning this research, I also recognised that violence will still occur, whether it be exclusion or gender-based violence. In this way, Hoffmann (2019) warns that “[...] rather than take responsibility for the ways we are daily and actively complicit in reifying culturally situated violences, we externalize bias and -after a few all-day seminars- count our demons exorcised”. From the beginning of these two studies, I knew the one-off interventions would not be a fix-all. Participants could still very well experience exclusion and gender-based violence. For me, I felt a duty to ensure the study’s participants and their safety and wellbeing, especially those with less societal power and privilege, are protected throughout the study. This meant ensuring participants know where to further seek support throughout and beyond the study itself. I felt this would be the case even if the interventions proved to be positively affecting the university climate.

For this reason, I knew that university resources needed to be embedded and made visible throughout the research process. For the gender-based violence study, I ensured that a resources list of university and non-university resources was attached to almost every email I sent to participants. The same resources list was found at the end of both surveys, and printed out for students to keep after the in-person focus groups. At the end of the study when the raffle prize winners were announced, the resources list was attached one final time. One student told me they were taking the resources list from the focus group to hang up in their hall’s communal kitchen for all of their housemates to see. Emails and face-to-face interactions for the inclusion study were less frequent, so I ensured a similar resources list was included in the surveys circulated too. Luckily, the university had ample resources available to participants, which made me feel more secure in mitigating ongoing violence for participants even as the research came to an end.

3.6: Chapter Conclusion

In this chapter, I provided an overview of the field site, data collection and data analysis for both the inclusion study and gender-based violence study that constitute this PhD research. I outlined how I use survey and thematic analysis methodologies to
complement more novel computational social science methodologies, which remain underutilised in the fields of inclusion and gender-based violence research. I specifically discussed ethical considerations and limitations that relate to conducting qualitative interviews, as well as ethical considerations and limitations more common in computational research. I also reflected on how my own biases may have influenced pieces of the research process. This chapter sets the foundation for the research design that will guide the following four chapters which focus on my findings.

**Chapter 4: Quantifying Trust with Mixed-Methods**

As discussed in Chapter 1, gender-based violence is just one aspect of the inequality regimes that operate within higher education institutions. However, both the literature and practitioners within universities are generally preoccupied with more overt manifestations and consequences of gender-based violence. This is necessary given the harm processes of gender-based violence cause at both the individual and group-level, but limiting. This chapter instead studies a more covert consequence and manifestation of gender-based violence at the group-level. Namely, this chapter will focus on perceptions of trust from dialogue and explore how group-level trust shifts before and after a gender-based violence sexual consent and active bystander training.

Trust may be defined as one’s confidence that an entity or another person will behave and have the outcome that is expected (Mohammadi & Hashemi Golpayegani, 2021). I will discuss three types of trust in this chapter including self-trust, social trust, and institutional trust. I refer to self-trust in the same way others like Banyard (2011) refer to self-efficacy, whereby one maintains trust in one’s skillset, in this case, to recognise gender-based violence and effectively intervene. Social trust is trust in others, which can be influenced by someone's individual traits, cognition, and personal preference (Fuglsang & Jagd, 2015). Important in and of itself, but also due to its effects on both self- and social trust, is institutional trust, or trust in one’s higher education institution. This type of trust is referred to as ‘thin’ trust as it usually lacks the strong sense of norms maintained in ‘thick’ trust between two people or more with a close relationship (Gerbasi & Cook, 2009).

With that, drawing on Banyard’s (2011) ecological framework, it is critical to discuss the relationship between these three forms of trust. Banyard (2011) discusses how the exosystem is made up of campus communities, including perhaps campus leaders, institutional centres, institutional protocols, and campus events. Altogether these institutional personas, structures, processes, and policies influence what the social norms may be on campus (Banyard, 2011). If a university has a highly visible and robust gender-based violence strategy in place, it may have GBV-related campus representatives, a multi-touchpoint GBV training programme for students and staff, a
rape crisis centre, posters for helplines in every bathroom or social centre, etc. All of this may affect the extent to which a student may trust their institution when it comes to both preventing and handling gender-based violence incidences. In turn, the exosystem’s social norms provide the context in which individuals navigate the microsystem. In the microsystem, students make sense of their relationships with others, and hence their trust in others too (Fuglsang & Jagd, 2015). In this way, if a student has high levels of institutional trust, perhaps they maintain higher levels of social and self-trust too. If the inverse is true and an institution lacks concrete GBV-related entities and policies, perhaps students will maintain lower levels of social and self-trust when it comes to both recognising and intervening in GBV incidents.

Regardless, trust, in all forms, relates to a variation in risk, uncertainty, and sense-making between two or more entities (Gerbasi and Cook, 2009). Gerbasi & Cook (2009) highlight that trust is tricky in that it is not just dyadic, but rather is part of these larger social networks, reminiscent of Banyard’s (2011) ecological framework for bystander intervention.

As previously mentioned, trust may be seen as a covert consequence and manifestation of gender-based violence as it may directly impact whether bystanders effectively intervene or not. For instance, a lack of trust in oneself in recognising gender-based violence or one’s intervening skills, means that one may be less likely to intervene (Banyard, 2008; Yule & Grych, 2020). This allows gender-based violence to persist. A lack of trust in other students intervening, or social trust, similarly reflects a context in which it is less likely that others will intervene if gender-based violence is witnessed. Distrust in one’s university, or institutional trust, may mean one is less likely to engage with university gender-based violence programming and mitigation initiatives. This once more allows gender-based violence to persist more easily. Therefore, while overt behavioural manifestations of gender-based violence are of course harmful, so too is a harder to see and quantify lack of self-trust, social trust, and institutional trust.

This chapter will hence unpack how group-level perceptions of trust shift before and after the gender-based violence active bystander training in terms of these three forms of trust. This work builds on Banyard’s (2011) ecological framework by investigating how the microsystem (social trust) and exosystem (institutional trust) feedback into preventing or upholding gender-based violence. In a way, social trust may be seen as bilateral trust within the microsystem, while institutional trust may be seen as bottom-up trust within the exosystem. This investigation of trust and multiple systems allows higher education institutions to understand how the inequality regime is being held in place, even after a sexual consent and active bystander training implementation. Accordingly, trust provides insight into who has power in a situation to intervene, thus reclaiming power from anyone seeking to make a claim on another’s bodily autonomy.
To accomplish the above, this chapter draws on focus groups and surveys collected to run a thematic analysis and survey analysis. I use thematic analysis to uncover the major themes from the focus group sessions. From there, I analyse key parts of the survey that have to do with bystander expectation and permission, community, safety, and trust. Gender-based violence promotes inequalities not just along gender divides, but also as a result of race, sexual orientation, disability, and more. Consequently, I also attempt to analyse how trust varies for different categorical groups and to understand how that impacts power relations within the university. Given the size of the sample, I conduct this analysis in terms of gender, sexual orientation, and race. Weaving the findings together, I discuss how trust, or lack thereof, may be allowing the university's inequality regime, and gender-based violence in particular, to persist. I also compare the findings to the literature. From there, I further add how the findings allow higher education institutions to adapt their training interventions to better address what is allowing gender-based violence to persist. I conclude this chapter by showing how investigating these more covert perceptions of (dis)trust at the group-level translates into the university needing to mandate training programmes, expand content, and partner with non-university entities to more effectively mitigate gender-based violence.

4.1: Data

I collected the data during the spring term 2022 through semi-structured focus group conversations and pre- and post-training surveys. I recorded the focus group conversations and then transcribed them using Otter.ai. I personally reviewed each transcription twice, and a research assistant reviewed them once. We preserved any ‘uh’, ‘like’, ‘um’, and pauses to capture tone. I also found that the inclusion of these utterances and pauses help communicate to readers how students sometimes stumbled speaking about these often-taboo topics. Six students completed pre-training focus groups, 11 students completed focus groups immediately post-training, and three students completed focus groups three weeks post-training. This split largely had to do with the fact that many students assigned to the three-weeks post-training focus group dropped out of the study. They either started ignoring emails, or dropped out due to workloads as the term progressed towards examinations. The focus group conversations were just about 1.5 hours in duration, but only a portion of the focus group data will be used in this section. In particular, the portion of the focus group that focused on trust and safety will be analysed with thematic analysis. This included such questions as “Would you feel safest living in your student accommodation with a person of your same gender- why or why not?” and “Would you lend your keys to other people in your student accommodation?”. I did not conduct thematic analysis to the parts of the
focus group conversation that focused on assessing knowledge of and attitudes towards gender-based violence as these parts were less about trust and more about general awareness of the concept of GBV. A total of 20 people participated in a focus group, or an online survey equivalent.

The portion of the focus group that involved the vignette activity will also be fully analysed in terms of trust. This activity included four vignettes being read out loud once for students. Names used in the fictionalised vignettes were pulled from a list of regionally common baby names. Students were asked to rank the four scenarios from most likely to intervene to least likely to intervene. After this, a semi-structured discussion about the four scenarios ensued. Follow-up questions focused on different aspects of the scenarios that impacted their ranking decisions and whether or not they would intervene. This included discussions on gender, sexual orientation, presence of alcohol, etc. Some of the key differences discussed can be seen in Table 4.1.

Vignette Scenario A: “One night, you’re out at a bar you frequent with your friends. It’s dark and crowded inside, but you stay close to your friend group. One of your friends, Isla, started drinking pretty early and it’s clear she is feeling pretty “good”. When she’s in this state, you know she gets flirty. Throughout the night, you see her repeatedly offering a boy she has met at the bar ample amounts of alcohol. She is making several flirty comments, all of which the boy is seemingly ignoring. She remains persistent in chatting with him.”

Vignette Scenario B: “It’s late one night and you’re walking home from the library. You’re trying to get home quickly since it’s dark out. The streets are practically empty. You’re pretty hungry so you decide to take the bus. As you wait at a bus stop, a guy walking by stops to talk to another girl waiting for the bus. It’s clear she is a student, and the guy says a generic pick-up line to try to engage in conversation with her. He seems sober, but keeps trying to talk to the girl despite the fact that she is uninterested. He asks if he can join her on the bus ride back to her place to get to know each other.”

Vignette Scenario C: “You’re going out for lunch and decide to go to the new and bright shopping centre to find some food. It’s crowded, but you manage to find a table at a café inside. At the table next to you, a girl is sitting down alone reading a chemistry textbook. A boy pulls up a chair at her table and introduces himself. The girl seems surprised and standoffish. He begins asking her questions about her book. The girl says she is trying to study, but he doesn’t get the hint. He keeps chatting her up.”

Vignette Scenario D: “You’re at a friend’s birthday party in a student accommodation. It’s a small, intimate gathering, and you know everyone in attendance. You see your friend Harris chatting to another friend Rory. You know Harris has a huge crush on Rory. Harris keeps asking Rory if he wants to take more shots of vodka. As Harris keeps drinking, he keeps touching Rory and mentioning that they should go back
to his place later to hang out more. Rory seems hesitant and keeps avoiding answering Harris’ advances.”

Table 4.1: Overview of Vignette Scenarios

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Scenario A</th>
<th>Scenario B</th>
<th>Scenario C</th>
<th>Scenario D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perpetrator Gender</td>
<td>Female</td>
<td>Male</td>
<td>Male</td>
<td>Male</td>
</tr>
<tr>
<td>Victim Gender</td>
<td>Male</td>
<td>Female</td>
<td>Female</td>
<td>Male</td>
</tr>
<tr>
<td>Bystander Gender</td>
<td>Unknown</td>
<td>Female</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>Perpetrator Sexual Orientation</td>
<td>Straight</td>
<td>Straight</td>
<td>Unknown</td>
<td>Gay</td>
</tr>
<tr>
<td>Location</td>
<td>Dark crowded bar</td>
<td>Bus stop at night</td>
<td>Shopping centre</td>
<td>Student accommodation</td>
</tr>
<tr>
<td>Touching</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>Yes</td>
</tr>
<tr>
<td>Presence of Alcohol</td>
<td>Present</td>
<td>None</td>
<td>None</td>
<td>Present</td>
</tr>
<tr>
<td>Bystander Awareness of Intent</td>
<td>Aware</td>
<td>Aware</td>
<td>Unknown</td>
<td>Aware</td>
</tr>
</tbody>
</table>

The Qualtrics survey completed by students occurred within one week pre-training, and then either within one week or three weeks post-training. This allowed for an understanding of student perceptions before and after the training, but also helped determine the longevity of the potential training effects on knowledge, skills, and awareness. 32 students completed the pre-training survey and 18 students completed the post-training survey. Just as with the focus group, not all survey questions evaluated trust. For that reason, only questions pertaining to trust, safety, and community will be analysed in this chapter. For the full list of questions in the survey, please see Appendix B.4. An example of a question that directly assessed trust was “I feel a sense of trust with other people in my student accommodation”. Indirect assessments of trust came through questions such as, “If I find myself a victim of unwanted sexual activity with a fellow student bystander present, I feel like I can expect the bystander to help me.” A total of 16 questions from the survey will be included in this chapter’s analysis. The questions were paired with a 5-point Likert scale (1= Completely disagree, 5= Completely agree).

To rule out potential changes in sample composition, it is important to discuss the differences between the samples for participants who completed Survey #1 and
Survey #2, as well as the pre- and post- focus groups. Overall, while the demographics of sample populations partly align with university demographics, it is important to emphasise that students voluntarily participated. For that reason, participants in the respective samples are likely to be a group of highly self-selecting students. Even so, for the survey samples, there were no major differences except an increase in non-heterosexual students, an increase in white-identifying students, and a decrease in students who previously completed consent training. Specifically, the proportion of non-heterosexual students in Survey #2’s sample increased from 56.25% to 72.22%. White participants increased from 68.75% to 83.33%. Given BAME and non-heterosexual students are more at risk of gender-based violence, it is important to check whether there were changes in self-reported victim-survivor status (Klein et al., 2022; Coulter et al., 2017). There was no major change in self-reported victim-survivor status as it only increased to 28.13% from 27.78%. The sample for Survey #2 also saw a drop in participants who previously completed consent training, from 28.13% to 5.55%. This is important to note as it means students who took Survey #2 are likely newer to gender-based violence concepts and likely completed the training with fresher eyes than participants who completed Survey #1.

There were more differences in the samples for pre-training focus groups and post-training focus groups. The post-training focus groups had a lower percentage of males (21.43% to 50.00%), which also meant the percentage of women increased to 64.29% in the post-training focus group. This is more representative of the university itself. Just as in the survey samples, the percentage of non-heterosexual participants increased from 33.33% to 78.57%, while the percentage of BAME participants increased from 16.67% to 28.57%. Again, BAME and non-heterosexual students are more at risk of gender-based violence, but this is not reflected in the post-training focus group sample (Klein et al., 2022; Coulter et al., 2017). Rather, the percentage of participants reporting they are victim-survivors dropped from 50.00% to 7.00%. Crucially, participants identifying as being unsure whether they are victim-survivors increased from 0.00% to 42.86% from pre-training to post-training. This may indicate that those who were in the post-training focus groups perhaps did not previously have deep knowledge of gender-based violence until completing the training. Moreover, the number of people unsure of whether they were victim-survivors may have increased as students learned more from the training about the spectrum of GBV. This is confirmed in the percentage of students in the post-training focus groups who did not have previous bystander training nor consent training. These percentages jumped from 66.67% pre-training to 92.86% post-training, and 66.67% pre-training to 85.71% post-training, respectively.

4.2: Analysis
4.2.1: Thematic Analysis of Focus Group Conversation

From the iterative coding process, five main themes emerged. I split the coding between pre-training and post-training focus groups to see conversations and themes pre-training and post-training. Pre-training focus groups saw 19 unique codes emerge, while post-training focus groups had 24 unique codes. Four themes related to self-trust and social trust and one theme involved institutional trust in mitigating gender-based violence. The four themes relating to self-trust and social trust were social relationship dynamics, gender dynamics, overall gender-based violence knowledge and awareness, as well as environmental factors. The gender-based violence knowledge and awareness theme emerged only in the post-training focus groups. Each of these themes will be discussed separately.

In the post-training focus groups, students overwhelmingly acknowledged that while their self-trust and social trust in others to intervene was highly situation dependent, they did feel more knowledgeable about gender-based violence. This newfound knowledge in turn fed into their trust in the student body intervening. With this new knowledge, students felt that they had increased confidence in their skillset to intervene, a better understanding of what is right and what is wrong, better knowledge of what escalation looks like, an enhanced understanding of what can be considered a violation to someone’s comfort levels, and an increased awareness of visual indicators of gender-based violence. This heightened sense of morality and determination to intervene appeared in such quotes as "even if I didn’t feel very comfortable, you know, the right thing to do, is the right thing to do" and "I’m certain that this is wrong behaviour, rather than, you know, I, I could be mistaken". It was clear that this newly gained knowledge helped students trust themselves and others to intervene. For example, students most often cited physical touch as an indicator of gender-based violence that would increase their levels of self-trust and social trust to recognise and intervene in any given bystander scenario. This is illustrated by one student reflecting, "[...] I can see physical contact, you know, and if I know that the person is not consenting, or is, is unable to consent, it’s, it’s a very clear sign that something is going wrong there. You know, I would doubt myself less". This shows how self-trust relates back to decreasing doubt in reading a situation as gender-based violence and, as a result, warranting bystander intervention. Students also specifically mentioned ways of intervening common in the literature and mentioned in the training itself, such as discreetly asking for help or creating fictitious situations. From the post-training focus groups, I was able to observe that students felt more self-trust and social trust in others to intervene when necessary, especially when it came to the training aiding in their ability to correctly read what is and is not gender-based violence.
The next theme that emerged in both the pre-training and post-training focus groups was the theme of social relationships dynamics’ impact on trust. In the pre-training and post-training focus groups alike, students expressed being able to trust their immediate contacts. Immediate contacts could be their close friends or acquaintances. In the pre-training focus group too, there was a very palpable social distrust in strangers, including students they did not personally know. One student remarked, “I feel like, like, intervening when it's your friends is a lot easier than intervening when it's strangers”, while another student reflected on a vignette saying, “everyone knows each other, like in that situation, so I feel like... they're less likely to get like, confrontational if everybody knows... And then there's other people there who like, would side with you, like if you intervened. They'd probably back you up.” Students in the pre-training group reflected on their self-trust intervening for friends and social trust in friends intervening on behalf of them. This remained true whether they were the perpetrator, victim, or bystander in a scenario. They felt friends were easier to read and that they could trust friends not to get violent. With strangers, on the other hand, this was less the case. Strangers represented wildcards and in scenarios where strangers are the victim or perpetrator, students felt less trusting of bystanders to intervene. There also seemed to be an idea that there is power in numbers when it comes to being in your own social circles, whereby everyone is keeping an eye out for each other and ready to back each other up.

Yet, in the post-training focus group, while students still spoke about trusting friends and acquaintances, students’ distrust in the overall university student body and community became more apparent. That does not mean there was absolutely no social trust in the general student body to intervene. Rather, students hinted at the idea that while the student body as a whole is friendly, that does not directly translate into trusting just anyone. For students who identified as sexual or gender minorities, this seemed more so the case too. These sentiments are exhibited in the following two quotes:

“[…] if I'm surrounded by people that I don't know, um, I would like to believe someone will help me, but I don't know how much I would trust someone else to help me at that point. Well... Um, I feel like, uh, from the people I've met, although everyone has been friendly, but I know when there is like another layer of, um, either violence or, sort of, harassment, I know that some people might not want to be involved. So, I am not extremely comfortable.”

“Not at all [confident]. And that's pessimistic but like, yeah, from my past experiences, no. It's not particularly great. Yeah, I mean, not just about like, so I'm, I'm bi as well. And like, like there's, there's a lot of like, um, homophobia I have experienced. Um, and like, there's
Reasons for this lack of trust in the general student body, excluding those who students have direct or indirect relationships with, seemed to boil down to potential social repercussions and anxiety, as well as potential power differentials. In terms of potential social repercussions and anxiety, students feared causing a potential scene, misreading a situation, and just a general apprehension to strangers approaching them or approaching strangers. The only people who did not feel this way in the focus groups self-proclaimed themselves to be very outspoken and confident extroverts. These feelings were seen in the pre-training focus group too, but were more pronounced in the post-training focus group. One student stated, “Like, they might know each other for some, whatever reason they might, I don’t know, they might be exes or whatever, they might know each other”. These feelings of potential social stigma were combined with fear of power differentials too. To demonstrate, a female student reflected, “What if, what if the boy has the other, other friends saying I’m too… um, be, I’m too… I’m stepping in other, someone else’s business? And what if, what if the boy, or the, what if the boy or the girl try to hit me or something”. It was easy to note students’ felt limitations to their increased levels of self-trust and social trust in intervening in situations where there was no pre-existing relationship. With no pre-existing relationship, students felt there was increased potential for social embarrassment, and increased potential for more violence, including physical, to occur due to the unknowns of the situation. One student pointedly even stated, “I do not indulge in matters that does not involve/concern me”. This shows how while students felt a boost in their self-trust and social trust due to increased knowledge and skills, there were stark limits to this boost when it came to potential GBV scenarios involving strangers, including students they did not know.

The next theme I found was related to gender dynamics, and the influence of gender of all parties involved on trust to intervene. Gender of the hypothetical victim, of the perpetrator, and the bystander were all discussed. In the pre-training focus group, students reflected mostly on feeling no difference in terms of whether they preferred to live with the opposite gender. In contrast, when it came to discussing gender in gender-based violence scenarios, students felt more trust with other students of the same gender. Women felt women would intervene for them, while some men expressed it was easier to intervene for other men or to interact with a male perpetrator. One student shared the nuance of gender differences between bystander and victim by offering, “[...] one of my housemates is quite like tall, rugby, very built. And, he was like, ‘I like can’t go and like, ask them like, ‘Oh, are you okay?’ Because I’m relatively intimidating. So it might
just scare them off or like, you know, make it worse.’ So that’s definitely something that makes it, I think, more difficult for men to intervene”. This shows how while self-trust is benefited by increased knowledge, awareness, and skills, it is limited by the intricacies of gender dynamics. Another expressed how the gender of the perpetrator influences their thinking saying, “But I do think gender plays quite a big part on the level of confidence and the willingness to intervene especially, um, if the, if the kind of person who’s harassing them is a guy or a girl. Plays quite a big part.” Therefore, I found that in the pre-training focus group, participants felt gender dynamics constrained their self-trust and social trust.

Gender dynamics between victim, bystander, and perpetrator came up even more frequently in the post-training focus groups. After the training, more students stated they would prefer to live with non-males, with one female student reflecting on preferring to live with the same gender, “Yes, I would, the main reason is that when I come back from a night out once I am inside, I can be confident I will be completely safe and not feel uncomfortable. For example, I trust my female flatmates much more to help me get to bed when I am intoxicated.” This shift hints at perhaps the training invoking a sense of fear and social distrust in the students who completed the training. This was also highlighted by students sharing how the gender narratives surrounding gender-based violence do influence who they trust to intervene. Specifically, students in the post-training focus group discussed how they felt more trust with women not just in terms of intervening, but also in terms of who they would intervene for. Ultimately, they felt women were less physically intimidating, had a naturally deeper understanding of gender-based violence, and a higher sensitivity to recognising it happening. One student even reflected on how they trust women perpetrators more than male perpetrators to recognise what they are doing is wrong and stop. Reflecting on Scenario A, a student said, “[...] another reason why I was like, it’s not a particularly threatening situation is because I feel... obviously girls can like sexually assault guys, sexually harass them, but I feel like girls can sort of understand it more, uh, like gender-based violence, harassment, because it happens to them on a daily basis”. Even though the female in Scenario A was actively committing gender-based violence, the student felt less inclined to intervene because the student trusted that the female perpetrator would recognise she was committing GBV and stop without intervention. This emphasises that students felt the gender of the victim, bystander, and perpetrator all play a role in self-trust and social trust when it comes to mitigating gender-based violence. Furthermore, the gender dynamics theme showcased what gender narratives remain prevalent within gender-based violence.

The biggest theme in the focus groups related to how different environmental factors influence students’ self-trust and social trust. Across both sets of focus groups, students specifically felt less trust in students intervening when they were alone, at
night, and when alcohol was consumed by anyone in the situation. Only one person in
the pre-training focus group saw alcohol as a value-add to trusting bystanders in
intervening because it may boost someone’s ‘liquid courage’. This was not the
majority’s viewpoint. Students largely felt that being alone, any of the parties consuming
alcohol, and it being nighttime added to the increased danger for the potential
bystander and the potential victim. Being alone means that there is no power in
numbers, with a student reflecting on Scenario B, “I, uh, think the the amount of people
that can do something in that situation is massively reduced compared to the other ones”.
Related to this, students highlighted the differences between social trust in the street
alone vs. in a crowded bar. Some reflected on the fact that bars mean it is easier to lose
track of others and to get distracted, whereby “[...] in a bar, that’s...I don’t know, it’s like
safer than outside, but it’s still like, if it’s dark and crowded, it’s very easy to lose people,
um, and very easy for someone to like, disappear. And then you don’t know where they’ve
gone or what they’re doing or who they’re with, so...”. Even so, students felt significantly
more self-trust and social trust in a bar scenario, but both being in the street alone at
night and in a bar at night prompted lower feelings of trust. This can be summarised by
concerns for safety and potential for distractions.

The final theme to emerge in the thematic analysis of focus groups was
institutional distrust. While this was a small theme in the focus groups, I felt it gravely
alarming that no students reflected on their trust in the university in mitigating gender-
based violence. Thus, while students felt at least to some extent self-trust and social
trust, this did not translate to institutional trust. Rather, students in the pre-training
focus group felt the organisation’s resources were pointless, too retroactive, suffered
from major time gaps in coverage and support, and retraumatising based on others’
experiences. A student spoke about hearing experiences with university counselling as
cold and feeling like an “interrogation”. They spoke of how the effort of going through
the institutional process did not feel worth the support ultimately received. These strong
sentiments of distrust were reflected in the post-training focus group too, but focused
more on how the university is not even the appropriate entity for support when it comes
to gender-based violence. While some reflected on their hesitancy to contact the police,
one student pointed out “I don’t think I would contact the University at all. I would contact
the police, I think, I, I think the university wouldn’t be the appropriate people to handle that.”
The extent of these sentiments are reflected in one student proposing that the
university make its resources “widely known and independent of the uni so they are more
trustworthy”. This shows how students felt like the university and trust were antithetical
when it comes to matters of gender-based violence. In this way, while self-trust and
social trust were evident at times, institutional trust severely lacked during both sets of
focus groups. Paradoxically, for trust in the university to exist, resources and support
needed to be completely separate from the university.
4.2.2: Survey Data Analysis

As discussed in Chapter 3, statistical analysis of survey data for the gender-based violence study was limited by the small sample size. A small sample size is usually defined as anywhere between n<15 through n<20 (Mircioiu & Atkinson, 2017; Norman, 2010). Given this, I was able to run the Mann-Whitney non-parametric test for my pre-training and post-training survey groups when post-training scores are aggregated. Unfortunately, my desire to run any additional simple modelling to understand the salience of gender, sexual orientation, or race and ethnicity in my outcomes is too risky. With such small sample sizes, most statisticians will agree that there is too much risk of losing external validity and running into Type II error whereby there is a false positive rejection of the null hypothesis (Norman, 2010). Consequently, for discussions of gender, sexual orientation, and race, I instead focus on discussions of the average scores pre-training and post-training. As a reminder, the Likert scale for these survey items was a 5-point Likert scale (1= Completely disagree, 5= Completely agree). Given the ordinal nature of the data (the difference between somewhat disagree to completely disagree cannot be assumed to be the same as the difference from somewhat agree to neither disagree nor agree), I discuss average scores as definitively higher or lower. Table 4.2, Table 4.4, and Table 4.6 present the in-depth statistical analysis of survey items pertaining to bystander expectation and permission, community, as well as safety and trust. Figure 4.3, Figure 4.5, and Figure 4.7 present some of these findings in bar chart format.

As seen in Table 4.2 and Figure 4.3, the average scores for the two questions in the survey that pertained to bystander expectation and permission increased from pre-training to post-training. When bystander expectation and permission is detected, this can be understood as indicative of an unspoken norm that students will intervene and help, without pushback, if another student is in danger. The increase in bystander expectation and permission measures from pre-training to post-training echoes the focus group sentiments, albeit neither had significant differences from pre-training to post-training though. However, when broken down into immediately post-training and 3-weeks post-training, Item #2 about whether there is a culture of unspoken permission for bystander intervention saw a lower average score of 3.4615 for the 3-weeks post-training group compared to the average score of 4.2000 for the immediately post-training group. This potentially shows that immediately post-training, students felt inspired by the training and as if there is a campus culture of unspoken permission for bystander intervention between students. As time progressed though, perhaps students reflected more and felt this was less the case, or the training simply had a short run influence on bystander permission. Another potential reason for this may be drawn from the thematic analysis. Students reflected a lot in the focus groups on their
increased knowledge, which added to their self-trust and social trust. As time went on though, students may have realised that this increased knowledge is limited to those students who took the training. As a result, trust in other student body members waned from immediately post-training. This would be aligned with the literature whereby trainings do not contribute to a lasting impact on bystander efficacy or willingness to intervene (Evans et al., 2019).
Table 4.2: Averages and Significance of Bystander Expectation & Permission Scores

<table>
<thead>
<tr>
<th>Statement</th>
<th>Item 1: If I find myself a victim of unwanted sexual activity with a fellow student bystander present, I feel like I can expect the bystander to help me.</th>
<th>Item 2: If I find myself a victim of unwanted sexual activity with a fellow student bystander present, I feel like there is a culture of unspoken permission for the bystander to help me.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Training Average (n=32)</td>
<td>2.9375</td>
<td>3.2188</td>
</tr>
<tr>
<td>Post-Training Average (n=18)</td>
<td>3.3889</td>
<td>3.6667</td>
</tr>
<tr>
<td>Delta Δ (Pre- to Post-)</td>
<td>+0.4514</td>
<td>+0.4479</td>
</tr>
<tr>
<td>Mann-Whitney P-Value</td>
<td>0.1779</td>
<td>0.1555</td>
</tr>
<tr>
<td>Immediately Post-Training Average (n=5)</td>
<td>3.2000</td>
<td>4.2000</td>
</tr>
<tr>
<td>Delta Δ (Pre- to Imm. Post)</td>
<td>+0.2625</td>
<td>+0.2427</td>
</tr>
<tr>
<td>3-Weeks Post-Training Average (n=13)</td>
<td>3.4615</td>
<td>3.4615</td>
</tr>
<tr>
<td>Delta Δ (Pre- to 3-Weeks Post)</td>
<td>+1.2625</td>
<td>+0.2427</td>
</tr>
<tr>
<td>Delta Δ (Imm. Post to 3-Weeks Post)</td>
<td>+0.2615</td>
<td>-0.7385</td>
</tr>
</tbody>
</table>

* Significant at p-value < 0.05
Figure 4.3: Likert Scale Responses and Averages per Survey for Bystander Expectation & Permission Items

Note that, where possible, I try to incorporate colour-blind friendly colour palettes into my figures.

* Significant at p-value < 0.05

4 Note that, where possible, I try to incorporate colour-blind friendly colour palettes into my figures.
For the modified community scale in Table 4.4 and Figure 4.5, all overall scores increased from pre-training to post-training except for one survey item. Even so, no differences proved statistically significant post-training. The one survey item that did not increase from pre- to post-training was Item #2- Community in Accommodation Importance. The average score for this survey item decreased from 3.7188 to 3.6667 from pre-training to post-training. A reason for this may be because post-training, students realised they feel a greater need for a sense of safety and trust, rather than community. Thus, their scores decreased post-training. When looking at differences from immediately post-training to 3-weeks post-training, exactly half of the survey items remained higher at 3-weeks out from the training. Two of these had to do with feeling a sense of community with the larger university community and in their student accommodations. This again may be since students realised they felt less of a need for community, and more of a need for safety and trust. Tying this to the thematic analysis does show that students do feel a sense of community to a certain extent with others at the university and in their student accommodations. As mentioned, participants did feel other students they met were friendly, nice, and mostly welcoming. This did not directly translate into having feelings of community per se.
### Table 4.4: Averages and Significance of Community Scale Scores

<table>
<thead>
<tr>
<th>Statement</th>
<th>Pre-Training Average (n=32)</th>
<th>Post-Training Average (n=18)</th>
<th>Delta Δ (Pre- to Post)</th>
<th>Mann-Whitney P-Value</th>
<th>Immediately Post-Training Average (n=5)</th>
<th>Delta Δ (Pre- to Imm. Post)</th>
<th>3-Weeks Post-Training Average (n=13)</th>
<th>Delta Δ (Pre- to 3-Weeks Post)</th>
<th>Delta Δ (Imm. Post to 3-Weeks Post)</th>
<th>Significance at p-value &lt; 0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1- I feel a sense of community with other people in my student accommodation (for example, you share interests and concerns with them).</td>
<td>2.9688</td>
<td>3.7188</td>
<td>3.8750</td>
<td>3.7813</td>
<td>4.1875</td>
<td>4.5313</td>
<td>3.8000</td>
<td>3.6000</td>
<td>4.2000</td>
<td>4.0000</td>
</tr>
<tr>
<td>Item 2- It is very important to me to feel a sense of community with people in my student accommodation.</td>
<td>3.1111</td>
<td>3.6667</td>
<td>4.2778</td>
<td>3.8333</td>
<td>4.3889</td>
<td>4.5556</td>
<td>0.7254</td>
<td>0.8420</td>
<td>0.1341</td>
<td>0.8359</td>
</tr>
<tr>
<td>Item 3- Some people care a lot about the kind of student accommodation they live in. For others, student accommodation is not important. What the student accommodation is like is very important to me.</td>
<td>+0.1422</td>
<td>-0.0521</td>
<td>+0.4028</td>
<td>+0.0520</td>
<td>+0.2014</td>
<td>+0.0243</td>
<td>-0.2188</td>
<td>+0.325</td>
<td>+0.2125</td>
<td>-0.3313</td>
</tr>
<tr>
<td>Item 4- I feel a sense of community with other people at my university (for example, you share interests and concerns with them).</td>
<td>0.7254</td>
<td>0.8420</td>
<td>0.1341</td>
<td>0.8359</td>
<td>0.3600</td>
<td>0.7177</td>
<td>3.8000</td>
<td>3.6000</td>
<td>4.2000</td>
<td>4.0000</td>
</tr>
<tr>
<td>Item 5- It is very important to me to feel a sense of community with people at my university.</td>
<td>+0.8312</td>
<td>-0.1188</td>
<td>+0.325</td>
<td>+0.2187</td>
<td>+0.2125</td>
<td>-0.3313</td>
<td>2.8462</td>
<td>3.6923</td>
<td>4.3077</td>
<td>3.7692</td>
</tr>
<tr>
<td>Item 6- Some people care a lot about the kind of university they go to. For others, the university is not important. What the university is like is very important to me.</td>
<td>-0.1226</td>
<td>-0.0265</td>
<td>+0.4327</td>
<td>-0.0121</td>
<td>+0.1971</td>
<td>0.161</td>
<td>-0.9538</td>
<td>+0.0923</td>
<td>+0.1077</td>
<td>-0.2308</td>
</tr>
</tbody>
</table>

* Significant at p-value < 0.05
Figure 4.5: Likert Scale Responses and Averages per Survey for Community Items

* Significant at p-value < 0.05
In terms of the safety and trust questions in Table 4.6 and Figure 4.7, all average scores increased from pre-training to post-training except for two survey items. These two survey items (Item #1 and Item #2) asked about feeling a sense of trust with people in student accommodations and while they did decrease post-training, they did so minimally. For feeling a sense of trust with fellow students in their accommodations, averages scores decreased from 3.5313 pre-training to 3.500 post-training. For the importance of feeling a sense of trust with student accommodation peers, scores decreased from 4.4063 pre-training to 4.3889 post-training. Results do align with what was discussed in the focus groups as post-training focus groups highlighted a strong desire to feel a sense of trust with overall peers at university. Item #8- Safety with Uni Peers Importance received an average score of 4.3438 pre-training and jumped to 4.8333 post-training (statistically significant at p= 0.0156). This confirms that students felt a heightened desire to feel a sense of safety with university peers post-training. Furthermore, it shows that students perhaps place more importance on overall student body dynamics than the dynamics within their student accommodations. Additionally and unfortunately, across the board for all trust and safety scores, average scores were lower three weeks post-training than immediately post-training. This further hints at the training perhaps having a lingering effect on making students realistically and critically reflect on their surroundings and relationships with the average university student. Simply, taking time to reflect on the training may have led to decreased feelings of safety and trust in the student body as a whole.
Table 4.6: Averages and Significance of Safety & Trust Scores

<table>
<thead>
<tr>
<th>Statement</th>
<th>Item 1- I feel a sense of trust with other people in my student accommodation.</th>
<th>Item 2- It is important to me to feel a sense of trust with the people in my student accommodation.</th>
<th>Item 3- I feel a sense of safety with other people in my student accommodation.</th>
<th>Item 4- It is important to me to feel a sense of safety with the people in my student accommodation.</th>
<th>Item 5- I feel a sense of trust with other people at my university.</th>
<th>Item 6- It is important to me to feel a sense of trust with the people at my university.</th>
<th>Item 7- I feel a sense of safety with other people at my university.</th>
<th>Item 8- It is important to me to feel a sense of safety with the people at my university.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Training Average (n=32)</td>
<td>3.5313</td>
<td>4.063</td>
<td>3.9375</td>
<td>4.5938</td>
<td>3.4063</td>
<td>4.2188</td>
<td>3.7500</td>
<td>4.3438</td>
</tr>
<tr>
<td>Post-Training Average (n=18)</td>
<td>3.5000</td>
<td>4.3889</td>
<td>4.0000</td>
<td>4.8333</td>
<td>3.7778</td>
<td>4.5000</td>
<td>3.8889</td>
<td>4.8333</td>
</tr>
<tr>
<td>Delta Δ (Pre- to Post)</td>
<td>-0.0313</td>
<td>-0.0174</td>
<td>+0.0625</td>
<td>+0.2395</td>
<td>+0.3715</td>
<td>+0.2812</td>
<td>+0.1389</td>
<td>+0.4895</td>
</tr>
<tr>
<td>Mann-Whitney P-Value</td>
<td>0.8983</td>
<td>0.4688</td>
<td>0.8119</td>
<td>0.0669</td>
<td>0.1252</td>
<td>0.4612</td>
<td>0.5820</td>
<td>0.0156*</td>
</tr>
<tr>
<td>Immediately Post- Training Average (n=5)</td>
<td>3.8000</td>
<td>4.0000</td>
<td>4.0000</td>
<td>5.0000</td>
<td>4.0000</td>
<td>4.6000</td>
<td>4.2000</td>
<td>5.0000</td>
</tr>
<tr>
<td>Delta Δ (Pre- to Imm. Post)</td>
<td>+0.2687</td>
<td>-0.0063</td>
<td>+0.4625</td>
<td>+0.4062</td>
<td>+0.5937</td>
<td>+0.3812</td>
<td>+0.4500</td>
<td>+0.6562</td>
</tr>
<tr>
<td>3-Weeks Post- Training Average (n=13)</td>
<td>3.3846</td>
<td>4.3846</td>
<td>3.8462</td>
<td>4.7692</td>
<td>3.6923</td>
<td>4.4615</td>
<td>3.7692</td>
<td>4.7692</td>
</tr>
<tr>
<td>Delta Δ (Pre- to 3-Weeks Post)</td>
<td>-0.1467</td>
<td>-0.0217</td>
<td>-0.0913</td>
<td>+0.1754</td>
<td>+0.2860</td>
<td>+0.2427</td>
<td>+0.0192</td>
<td>+0.4254</td>
</tr>
<tr>
<td>Delta Δ (Imm. Post to 3-Weeks Post)</td>
<td>-0.4154</td>
<td>-0.0154</td>
<td>-0.5538</td>
<td>-0.2308</td>
<td>-0.3077</td>
<td>-0.1385</td>
<td>-0.4308</td>
<td>-0.2308</td>
</tr>
</tbody>
</table>

* Significant at p-value < 0.05
Figure 4.7: Likert Scale Responses and Averages per Survey for Safety & Trust Items

* Significant at p-value < 0.05
Gender seemed to also play a role in how the training affected students, with the breakdown of demographics per survey in Table 4.8. Additionally, in Table 4.9, for the bystander expectation and permission scale when average scores were aggregated, females and gender minorities saw a lower average score pre-training than males. The largest difference between the pre-training scores on the scale were between males and gender minorities, with gender minorities averaging 5.00/10.00 and males averaging 6.75/10.00. There were smaller differences between genders on the pre-training scores for the community scale, as well as the safety and trust scale. Interestingly, gender minorities and females actually had higher summative community scores than males pre-training as in Table 4.10. This did reverse post-training with males jumping to an average cumulative score of 25.25/30.00, compared to females’ 23.09 and gender minorities’ average cumulative score of 24.67. Post-training, the trust and safety scale average scores had no big differences between genders. They were all within +/-1.00 of each other for males, females, and gender minorities. Yet, what is interesting to look at is that the average score for gender minorities was the same pre-training and post-training. Males’ average score jumped from 32.00/40.00 to 37.25/40.00, while gender minority students’ cumulative score remained at 33.00. This potentially links back to the thematic analysis whereby men may feel a higher sense of trust and safety due to their (usually) larger physical presence and ability to intervene in dangerous situations.

In terms of sexual orientation, the only large difference between pre-training average scores was in the community scale questions. Students identifying as straight scored on average 23.71/30.00, while students part of the sexual minority scored on average 22.56/30.00. This potentially points to the fact that students who identify as straight may have a subtly easier time finding community during their first year of university, or have tighter knit communities. Post-training, the average cumulative score for sexual minority students was lower at 22.15/30.00. This was 28.20/30.00 for straight students which potentially shows a higher desire for having community than pre-training. A similar difference was seen for straight students on the trust and safety scale in Table 4.11. This may be because straight students felt more prepared from the training to intervene and had higher trust in their own knowledge and skills than students who identified as not straight. For instance, there is a chance non-straight students felt less prepared and trusting in the knowledge they gained as it pertained to queer gender-based violence situations.

Lastly, in terms of racial differences in the survey data, similar sentiments towards community were seen as with straight and non-straight students. The average cumulative scores on the community scale were actually higher for BAME students than for white students pre-training. Again, this may allude to tighter knit minority communities in a predominantly white university. On the trust and safety scale, as well as the bystander expectation and permission scale, both white students and BAME
students had nearly the same post-training scores of 33.80 and 33.33 of 40, respectively. White students’ scores were only marginally higher for both of these scales. Trust and safety was higher by a full +2.93 for BAME students in terms of trust and safety post-training. Taking all of this together shows that perhaps the training helped alleviate differences between white students and BAME students, but that the training only exacerbated differences between straight students and non-straight students. The same could be said when it comes to the differences between males and gender minorities in terms of trust and safety too.
### Table 4.8: Demographics of Survey Respondents

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
<th>Gender Minority</th>
<th>Straight</th>
<th>Sexual Minority</th>
<th>White</th>
<th>BAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Training (n)</td>
<td>20</td>
<td>8</td>
<td>4</td>
<td>14</td>
<td>18</td>
<td>22</td>
<td>10</td>
</tr>
<tr>
<td>Post-Training (n)</td>
<td>11</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>13</td>
<td>15</td>
<td>3</td>
</tr>
</tbody>
</table>

### Table 4.9: Bystander Expectation & Permission Scale by Demographic Group

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
<th>Gender Minority</th>
<th>Straight</th>
<th>Non-Straight</th>
<th>White</th>
<th>BAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Training Average</td>
<td>6.15</td>
<td>6.75</td>
<td>5.00</td>
<td>6.29</td>
<td>6.06</td>
<td>6.32</td>
<td>5.8</td>
</tr>
<tr>
<td>Post-Training Average</td>
<td>7.00</td>
<td>7.05</td>
<td>6.67</td>
<td>6.80</td>
<td>7.15</td>
<td>7.07</td>
<td>7.00</td>
</tr>
<tr>
<td>Delta Δ (Pre- to Post)</td>
<td>+0.85</td>
<td>+0.30</td>
<td>+1.67</td>
<td>+0.51</td>
<td>+1.09</td>
<td>+0.75</td>
<td>+1.20</td>
</tr>
</tbody>
</table>

### Table 4.10: Community Scale by Demographic Group

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
<th>Gender Minority</th>
<th>Straight</th>
<th>Sexual Minority</th>
<th>White</th>
<th>BAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Training Average</td>
<td>23.40</td>
<td>21.88</td>
<td>23.75</td>
<td>23.71</td>
<td>22.56</td>
<td>22.82</td>
<td>23.60</td>
</tr>
<tr>
<td>Post-Training Average</td>
<td>23.09</td>
<td>25.25</td>
<td>24.67</td>
<td>28.20</td>
<td>22.15</td>
<td>24.07</td>
<td>22.67</td>
</tr>
<tr>
<td>Delta Δ (Pre- to Post)</td>
<td>-0.31</td>
<td>+3.37</td>
<td>+0.92</td>
<td>+4.49</td>
<td>-0.41</td>
<td>+1.25</td>
<td>-0.93</td>
</tr>
</tbody>
</table>

### Table 4.11: Trust & Safety Scale by Demographic Group

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
<th>Gender Minority</th>
<th>Straight</th>
<th>Sexual Minority</th>
<th>White</th>
<th>BAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Training Average</td>
<td>32.10</td>
<td>32.00</td>
<td>33.00</td>
<td>32.14</td>
<td>32.22</td>
<td>33.00</td>
<td>30.40</td>
</tr>
<tr>
<td>Post-Training Average</td>
<td>32.63</td>
<td>37.25</td>
<td>33.00</td>
<td>36.60</td>
<td>32.62</td>
<td>33.80</td>
<td>33.33</td>
</tr>
<tr>
<td>Delta Δ (Pre- to Post)</td>
<td>+0.53</td>
<td>+5.25</td>
<td>0.00</td>
<td>+4.46</td>
<td>+0.40</td>
<td>+0.80</td>
<td>+2.93</td>
</tr>
</tbody>
</table>
This analysis highlights the importance of investigating covert group-level sentiments to uncover nuances in gender-based violence research. Overall, the survey analysis, while limited in its statistical power, points to a few nuanced patterns that echo some focus group narratives. Notably, the average scores pre-training compared to post-training for the bystander expectation and permission items were higher, but lower again by three weeks post-training. Additionally, the survey hinted at students feeling a greater desire for safety and trust than community in their student accommodations. This was reflected in the focus groups and the significantly lower difference between post-training and pre-training survey results for Item #8 - “It is important to me to feel a sense of safety with the people at my university”. Interestingly, student sentiments of trust and safety were lower three weeks after the training compared to immediately post-training for all survey items. These findings align with the literature in that the lasting effects of sexual consent and bystander trainings is a mixed bag (Evans et al., 2019). Finally, in terms of survey analysis by demographic groups, gender was found to be the most salient identity for differences in the average cumulative bystander expectation and permission scores, as well as the trust and safety average scores. At the same time, it is noteworthy that sexual minorities and BAME students had lower average community scores post-training than pre-training. Sexual minorities also scored the lowest on average in the trust and safety scale. This once again showcases the pattern in the focus groups whereby males felt inherently safest due to their statures, but sexual minority and gender minority students hinted at being more unsure of their safety.

4.3: Discussion

This chapter showcases the importance of more covert group-level indicators of gender-based violence. Higher education institutions currently mostly focus on the group-level overt behaviours, such as the annual reports of acts of gender-based violence. While important, higher education institutions should also investigate perceptions that affect gender-based violence like group-level trust. Trust in this case encompasses self-trust, social trust, and institutional trust, as each affects prevention of gender-based violence. When investigated, these three indicators of gender-based violence present a number of key learnings and implications for higher education institutions hoping to combat gender-based violence within their student bodies. Accordingly, this chapter’s theoretical contribution lies in its showcasing of the covert and group-level investigation of trust as it relates to gender-based violence.

In terms of an empirical contribution, five main themes emerged from the thematic analysis while an additional analysis of the student surveys reinforced some themes. The four main themes pertaining to students’ self-trust and social trust were
social relationship dynamics, gender dynamics, gender-based violence knowledge and awareness, and environmental factors. Gender-based violence knowledge and awareness post-training was apparent, highlighting that post-training, students felt a boost in trust in their own and others’ intervention skill sets. This is critical especially for female potential bystanders as women are more likely than men to decide not to intervene because of a perceived lack of skills (Yule & Grych, 2020). The fifth theme highlighted students’ glaring distrust in the organisation itself in preventing and managing gender-based violence incidents. While students shared that their trust in other students is always going to be situation dependent, institutional distrust was relatively solidified regardless of the situation. This is important because institutional trust provides the social context in which students make sense of social trust (Fuglsang & Jagd, 2015). This poses the question of if students have a firm distrust in the institution, how is social trust inherently limited by this?

One of the main themes with major implications for higher education institutions combating gender-based violence was the relevance of social dynamics. Students reflected on seeing physical touch as sparking self-trust and social trust in others to effectively intervene, regardless of whether the victim and/or the perpetrator were strangers or friends. Students overall felt that they would always trust in friends and immediate contacts more to intervene than strangers. They felt the same in terms of when they felt most trusting of themselves to intervene too. This is a major learning for higher education institutions given that this translated to an overall lack of social trust in the general student body. Furthermore, this finding does align with the literature as it relates to friendships and gender-based violence intervention. Katz & Moore (2013) in a meta-analysis found that when the victim or perpetrator is a friend, students find it is significantly easier to intervene. Seo et al. (2022) in a study of U.S. university students found that potential bystanders are more likely to witness friends in gender-based violence scenarios. Furthermore, students are more likely to intervene for friends than strangers (Seo et al., 2022). Simply put, strangers are wild cards. They could be dangerous, as well as verbally and physically aggressive. Students agreed that friends are less likely to respond aggressively and generally one can guess how they will respond. Rather, there is more emotional weight to intervening with and on behalf of friends. This is also echoed by Seo et al. (2022) who found anxiety levels are higher when students witness friends in tricky situations rather than strangers. In a Blayney et al. (2021) study conducting focus groups, 83% of focus groups felt some sort of hesitancy in intervening on behalf of friends due to potential social consequences. For Yule & Grych (2020), first year university students cited fear of social repercussions as the least common barrier to intervening. This is more aligned with this chapter’s study, whereby even with the social risks, intervening for friends still seemed preferable to
strangers. This is because strangers seemed to automatically equate to potential danger to oneself and potentially more harm to the victim.

To build trust within the general student body, higher education institutions must seek to build and maintain an unspoken norm of bystander intervention in gender-based violence scenarios. Based on a Banyard et al. (2021) study, perceived peer helping and perceptions of community are both correlated with students intervening for friends and strangers alike. This points to universities having to address feelings of community for students. From the pre-training survey, BAME students and gender minorities reported the highest feelings of community. Importantly, post-training, the straight, white, and male students all reported the highest perceptions of community. Given statistical analysis could not be calculated for demographic groups, this study would need to be repeated with a larger sample size to investigate whether the training itself contributed in some way to making the historically marginalised groups realise that they lack community at university. What is interesting too is that Yule & Grych (2020) found in their study of first year university students, men are more likely to decide not to intervene as they tend to view gender-based violence scenarios as not their responsibility. Therefore, it would also be important to repeat this study with a larger sample size to see if male students continue to have the highest perceptions of community. This is critical to investigate whether community does or does not directly translate into men actively pursuing bystander intervention.

If certain demographics of students are failing to feel as integral to the community, this perhaps contributes to lower social trust between general student body members, which in turn helps gender-based violence persist. As a result, if universities wish to combat gender-based violence, it may be worthwhile to ensure all students feel similar levels of community. With that, universities must expand students’ social circles, not in terms of how many friends a student has, but in terms of making students feel more connected to the average peer. Although Seo et al. (2022) found students are more likely to witness incidents involving friends, that does not mean students do not witness gender-based violence involving students they do not know. With higher perceptions of community, students can feel power in numbers with every fellow student they pass regardless if they know them personally or not. As found in the survey, students post-training felt a higher sense of importance for feeling a sense of safety with fellow peers than students pre-training (p= 0.0156). One key way to address this is to mandate bystander intervention training for all students. Additionally, adding on an in-person segment with small, mixed-identity groups at the beginning of university can help students feel connected to other identity groups they may not necessarily interact with or have honest conversations with. Furthermore, this would set the tone for the rest of their time at university, ensuring students know that there is an unspoken norm of, regardless of which friend group or identity group you fall into, you can trust
other students will have your back. Once again, this shows how universities can cover all their bases by also incorporating strategies that acknowledge the importance of more covert group-level perceptions of gender-based violence.

The university and other higher education institutions may also finesse their training programmes based on the learning from the gender dynamics themes in this study. The findings showed that narratives of the male perpetrator and female victim prevail, even in the minds of students who identify as gender minorities. The effect of this shows a decreased amount of self-trust and social trust in intervening when scenarios do not depict these gender-based violence norms. When a scenario reflects a male perpetrator and female perpetrator, students find these situations easier to read and trust intervention will occur more frequently and quicker. This is aligned with the higher rates of perpetration by males than females (Walsh et al., 2019). Where students lose trust is when intervention occurs for opposite genders and non-gender normative scenarios. A repetition of the study’s survey would confirm whether or not males and straight students continue to outscore female students, students identifying as part of a gender minority, and sexual minority students. The gender dynamic theme also builds on the literature whereby Ermer et al. (2021) found a gendered double standard. While the participants in the study ranged from 18 years old to 70 years old, a trend was that men should not be aggressive as they could cause more harm than a woman. Females when aggressive though were viewed as minor threats in comparison, warranting less intervention by potential bystanders (Ermer et al., 2021). This research builds on these findings by showing how the dominant gender narratives of gender-based violence prevail more specifically in university students’ minds too.

For this reason, universities must expand the scenarios and narratives they use to teach students how to be an active and effective bystander to gender-based violence. This is very uncommon in bystander trainings, especially in terms of content pertaining to queer and gender minorities (Kirk-Provencher et al., 2021). Of course, this is despite these individuals experiencing gender-based violence at higher rates than their majority counterparts (Klein et al., 2022). Thus, higher education institutions should address specific questions in their training programmes. Some of these questions could include how can men intervene for women without coming across as creepy or domineering? How can women intervene for men with the same ease as intervening for other women? Furthermore, how can students learn to read situations involving non-heteronormative and gender minorities as confidently as they do the heteronormative female victim and male perpetrator narrative? One way higher education institutions can begin to do this is by incorporating discussions of power into the training programme itself or a small in-person group discussion afterwards (Flecha et al., 2020; Sulley et al., 2020). Even if universities feel they lack the ability to provide students with a list of tangible steps and ways to intervene in these types of scenarios, students can generate the knowledge
themselves. For instance, smaller in-person groups with mixed identities could come together to brainstorm how they would prefer to be intervened for. Universities would be handing over the power of generating knowledge to the people who are actually affected in gender-based violence scenarios. This may in turn help students better trust not just other students, but also their institution itself.

Scenarios whereby students were alone, on the street, in the dark, at a bar, or consuming alcohol also proved to decrease self-trust and social trust to intervene. This is minus one student who felt alcohol gave themselves and others ‘liquid courage’ to intervene. Regardless, given many students likely often find themselves in bars, consuming alcohol, or walking at night alone, the fact that students remain unsure in these situations is a key weakness of the training implemented. Students did feel that bars presented an opportunity to have power in numbers, whereas being alone on the street at night felt extremely dangerous comparatively. The literature actually does not touch on these types of scenarios, but focuses more on bar scenarios and scenarios with alcohol. In Blayney et al.’s (2021) study, 100% of students felt alcohol took away from their ability to intervene, while Walsh et al. (2019) found binge drinking by university students comes with higher odds of both ambiguous consent and sexual assault perpetration. Furthermore, students do feel “being sure” is always a point of hesitance, even without alcohol having been consumed (Brodt et al., 2021, p. 13; Foubert et al., 2010; Hillman, 2021). Likewise, the idea that bar scenarios presented difficulties for students to keep tabs on others also was explored in the literature. In fact, students often feel that while bars present power in numbers, it is balanced by the number of distractions including dancing and socialising (Blayney et al., 2021).

Just as with the issues of distrust arising from gender dynamics, higher education institutions should address the multiple critical barriers to building trust when it comes to environmental factors and beyond. This suggestion is aligned with Yule & Grych (2020) who found first year university students’ decision making of whether to intervene or not is usually impacted by multiple reasons to and not to intervene at once. Simply, the decision making is complex given gender-based violence scenarios are often complex. This is an easy lift on the university’s part as it would involve incorporating more scenarios into the training where the student would be alone, in the dark, and consuming alcohol. Of course, this would also need to address the intricacies of alcohol consumption. For example, a scenario where the potential bystander has consumed alcohol is not the same as when a perpetrator is clearly belligerent. Scenarios and resulting actions to effectively intervene will all differ depending on who in the scenario is drinking. Again, the university does not need to come to students with exact steps they should be taking in these cases. Rather, universities could allow students to workshop these types of scenarios together to determine what they would like the norm to be for intervening in those tricky scenarios.
Similarly, the training programme could also include conversations with off-campus organisations such as bar staff from student-frequented bars, or even bus drivers from student-frequented bus routes. The scenario that presented the most hesitancy for students was Scenario B where the student is at a bus stop alone with the victim and the perpetrator. Only one student addressed the fact that they would not be alone the moment the bus driver drove up to the bus stop. Under the university’s gender-based violence mitigation programme, a group of bus drivers, students, and bar staff could come together to workshop a secret word or hand signal students can give to indicate they need immediate and carefully considered help. In some bars in Britain, this already exists whereby consumers can “Ask for Angela”. A student-specific term may prove helpful too. Furthermore, a hand signal students can give to bus drivers who may not even be stopping at their bus stop could also prove useful in dire scenarios. This way, students can build a sense of trust with not just other students at the university, but also trust with the university itself as it expands its gender-based violence mitigation programmes.

Two paradoxes emerged in this chapter. The student participants reflected on having no trust in their higher education institution and felt to gain institutional trust, the university needed to completely remove itself from the steps it is taking to mitigate and manage gender-based violence. At the same time, the whole reason university staff commissioned this research is because they wished to expand their gender-based violence mitigation and management strategies. The other paradox my research highlights is that the sexual consent and active bystander training implemented targets the individual by trying to encourage individuals to step up in gender-based violence scenarios. Yet, the fact that there was no institutional trust shows that more research into and interventions catering to the larger social context are critical. To date though, interventions targeting the institutional or larger group-level remain largely ignored in gender-based violence research (O’Connor et al., 2021).

One reason for this is perhaps because higher education institutions, as neoliberal organisations, are becoming increasingly preoccupied with performativity and individual responsibility (O’Connor et al., 2021). As higher education institutions incorporate more market-oriented ideologies, they both knowingly and unknowingly turn to airbrushing issues within academia to preserve their public image (Phipps, 2018). This is a fragmented approach to gender-based violence as the culture of a university influences how gender-based violence scenarios play out, and in turn, what behaviours potential bystanders employ (O’Connor et al., 2021). This research shows how if interventions, such as trainings, are implemented and target individuals, but a glaring distrust of the institution remains, gender-based violence can still persist. For that reason, HEIs must stop engaging in blissful ignorance of more group-level issues such as institutional distrust because it directly allows harm to still be perpetuated by
individuals (Phipps, 2018). Furthermore, by ignoring and failing to solve institutional distrust, it allows higher education institutions to continue to avoid accountability. In this way, the training can perhaps be seen as part of the airbrushing that higher education institutions undergo, covering up the larger and more deeply rooted issue of complete institutional distrust.

Of course, while students suggested the university separate completely from gender-based violence mitigation strategies, no higher education institution in today’s climate is going to discontinue all support services. There is too much at stake. What universities can do is seek to build more self-trust, social trust, and build more extra-university relationships that in turn foster student trust in the university itself. By making bystander intervention training mandatory, the university can directly ensure students have a common understanding of what the unspoken norm is when it comes to trusting other students to be effective bystanders. If the university tailors these training programmes to address some of the remaining critical barriers students feel in trusting themselves and others in intervening, then the university will only reinforce the unspoken norm. This will return power to students who may potentially be at more risk than others, such as gender and sexual minority individuals. Given the difficulties in retaining participants for this study, it is abundantly clear that students are unlikely to complete a 3.5-5.0 hour training in their own time. The only way all students will complete the training is if the university mandates it. As part of this programme, students can also work with organisations external to the university itself, but still under the university’s programme umbrella. Lastly, this research shows how researchers at higher education institutions must continue to investigate group-level perceptions of trust that influence gender-based violence. Only then can HEIs pinpoint how interventions may be better aligned to enhance their mitigation of gender-based violence.

4.4: Limitations

A few key limitations should be discussed as they pertain to this chapter. Note, for a larger discussion of limitations of the study as a whole, please see Chapter 3. More specific to this chapter, limitations may be divided into limitations of data collection and of data analysis. The main limitation to the data collection itself regards the limited sample size and drop-off of student participants post-training. This is true for the focus groups and surveys alike. The focus groups had six students in the pre-training sessions, 11 students in the immediately post-training sessions, and just three students in the sessions three weeks post-training. The surveys had 32 responses pre-training, 5 responses immediately post-training, and 13 responses three-weeks post-training. The small sample sizes impacted the data analysis in that only the Mann-
Whitney U test could reliably be run on the combined post-training survey data to test for statistical significance without posing risk to losing external validity and encountering Type II error. The samples were also dependent and varied in some ways from pre-training to post-training. Dependent samples pose the risk that participants who took both surveys grew bored with the survey and their boredom influenced their responses in some way. In terms of limitations to the sample, it is worth noting that this study was well-resourced, especially for students in Cohort 2. The issues I encountered regarding sample sizes are likely to be encountered by any university wishing to replicate this work without mandating the training itself. This further highlights that the students were very likely to be a highly self-selecting group, given the voluntary basis of this study.

Another data collection limitation could be that students in focus groups influenced each other’s answers as a result of conformity bias, which contributes to groupthink. This was less likely in student focus group surveys, which students who felt uncomfortable in group settings completed in lieu of an in-person focus group. Related to this, it was clear that differences in knowledge levels were apparent when some students emerged as predominant voices commanding expertise of gender-based violence vocabulary. Other students clearly stumbled with the relevant vocabulary, taking time to learn how to talk about perpetrators, potential victims, victim-survivors, etc. Crucially, students with less knowledge did not seem to outwardly change their immediate responses based on other students’ opinions. I only took note of this happening once when one student convinced another that a scenario was more dire than another. The student had already written down their ranking of the scenarios on their piece of paper and handed it in though.

Limitations to data analysis were also apparent to an extent. Coding will always be subjective, but I tried to ensure robustness of my qualitative work through a systematic and iterative approach to coding. I also took some time between reviewing codes and their subsequent themes to remove myself from the data. Upon returning to the themes, I thoughtfully considered alternative interpretations while engaging in the careful reading. I’d like to note that by the time I coded the focus groups, I had already carried out previous coding and qualitative analysis with an undergraduate research assistant for another report on the same data. While this may have biased my thematic analysis coding, it also meant I was highly familiarised with the transcripts and contexts of what was being said already. Finally, quotes included in this chapter are my attempts at representing the themes. Of course, there are nuances to each theme that cannot be fully captured in any one quote. I do feel this chapter is an accurate representation of student sentiments and because of this, I hope to share my chapter with the student participants who asked for the study’s findings to be shared in the future.
4.5: Chapter Conclusion

Exploring covert group-level perceptions influencing gender-based violence as it relates to trust provides insight for universities hoping to mitigate and manage gender-based violence. Using thematic analysis and a survey analysis, this chapter explores self-trust, social trust, and institutional trust before and after a sexual consent and active bystander training. I carried out thematic analysis on the data from focus groups to find five major themes including: social relationship dynamics, gender dynamics, gender-based violence knowledge and awareness, and environmental factors. The fifth theme signified a complete student distrust of their institution. Major implications based on the learnings from these themes showed that the higher education institution could implement several strategies to foster trust between students and between student and institution. Such strategies could include mandating the training to build an unspoken norm for all students, regardless of if they are friends or strangers. Implementing small group cross-identity discussions after the training may help expand social circles beyond identity groups. Furthermore, expanding university programmes to include non-university entities such as bars and public transport services may further help create an unspoken norm of bystander intervention, as well as boost institutional trust. The chapter presents a noteworthy exploration of more covert and group-level sentiments of trust to evaluate a GBV intervention training programme and create a feedback loop for finetuning future higher education interventions.

Chapter 5: Quantifying Trust with Computational Text Analysis

There are three main goals of this chapter, which I will discuss in terms of a theoretical contribution, a methodological contribution, and an empirical contribution. Theoretically, I once again showcase how applying the new theoretical framework of inclusion/exclusion and gender-based violence is critical in unveiling more group-level and covert inequality regimes. More specifically, after the administration of sexual consent and bystander training, I evaluate covert sentiments of trust at the group-level. In doing so though, I look more directly at the participants’ behaviour, specifically their dialogue, to gauge their sentiments of trust. I show how ignoring more community-oriented factors that may influence the effectiveness of bystander intervention is failing the gender-based violence literature. Without researching the group-level and more covert aspects of inequality regimes, gender-based violence is inadvertently allowed to persist. Namely, (dis)trust between students at the microsystem level and (dis)trust between student and institution at the exosystem level contributes to potential inequalities of power on campus.
Methodologically, this chapter contributes to the literature by developing a method and illustrating how this method can be used on common gender-based violence text corpora. I exhibit how a computational text analysis method can draw the same conclusions as qualitative work, while, at the same time, shedding new light on the findings. Thus, while addressing the same research question as Chapter 4, this chapter’s computational text analysis presents itself as an equally insightful and more reproducible alternative to the qualitative methods I previously used. Rather than looking at the content of what is said as with thematic analysis, the text data studied is first rendered intelligible to later generate insight (Benoit, 2020). Accordingly, this chapter first develops a bespoke lexicon of trust words pulled from the most used gender-based violence trust survey items. This process is similar to, but more simplified than, Alsaid et al. (2022) who also created a unique trust lexicon. Given Alsaid et al.’s (2022) focused more broadly on trust questionnaires for ecommerce, automation, and human relationships, they were able to utilise word embeddings within their larger questionnaire text corpus to create their bespoke trust lexicon. Gender-based violence questionnaires are fewer in number, and therefore I simplified the process and decided to still use word embeddings. I used pre-trained word embeddings in lieu of in-text word embeddings. From there, this chapter applies the new dictionary of trust words to sentiment analysis, and compares the sentiment analysis to analysis done with two other dictionaries: the widely used off-the-shelf NRC dictionary and Alsaid et al.’s (2022) own bespoke trust lexicon. These two dictionaries present a different, but similar, list of trust-related words as they were not developed in the specific domain of gender-based violence. Through data-analysis triangulation and grounding the results in the qualitative findings of Chapter 4, I conclude that domain-specific lexicons such as mine and Alsaid et al.’s (2022) dictionary offer an underutilised, yet worthwhile, method for gender-based violence research.

Empirically, this study expands on the findings from Chapter 4. Using the bespoke trust lexicon, I find that social trust is higher than institutional trust once again. However, one of the added benefits of computational text analysis is that in addition to being more reproducible, it offers a numerical value of trust for comparison purposes. I conclude that if both social trust and institutional trust are low, the university should try to boost its institutional trust to increase social trust. Furthermore, this study is the first of its kind, to my knowledge, that quantifies institutional trust in comparison to social trust on a university campus. Keeping these findings in mind, I emphasise that group-level and covert (dis)trust of students in the institution is a conclusion that once again cannot be ignored by researchers. I finish this chapter discussing how further research into the heart of the reasons for why students feel low levels of institutional trust is needed. I also discuss how although my datasets are limited in size, my chapter provides a step-by-step guide to applying computational text analysis to gender-based
violence research. With that, sexual consent and active bystander trainings should be made mandatory or more incentivised for university students. This will not only address campus culture, but also help generate a larger corpus in the future to further explore more group-level and covert mechanisms influencing gender-based violence.

5.1: Data

As discussed in Chapter 3, the data used in this study were collected through semi-structured focus groups and semi-structured individual interviews. Each of these were recorded, and then transcribed with the help of Otter.ai. I then manually reviewed the transcriptions twice more, and a research assistant reviewed them once too. While 20 students completed focus groups, only the post-training focus groups were used in this chapter’s analysis. Thus, the focus group data incorporated data from 14 students, of which only three students were in the three weeks post-training focus group. The remaining 11 students all completed the focus group within one week of their training. All focus groups except for one group of two students were carried out in-person. This is because one of the students had already returned home for the summer break. The full focus group was about 1.5 hours long and occurred in either late spring 2022 or autumn 2022. For the purpose of this chapter, only certain questions from the focus group will be analysed. Given this chapter is concerned with students’ perceptions of social trust and institutional trust, as defined in Chapter 4, I selected questions from the focus group that pertain to these concepts. As the focus groups were also to analyse students’ change in knowledge and awareness of gender-based violence, many of the focus group questions did not specifically pertain to trust. The three questions analysed from the focus groups included two social trust (student-to-student) related trust questions, as well as one institutional trust (student-to-institution) trust question.

The 1:1 interview data used in this chapter were collected in early summer 2022, just after exam season. A total of seven students participated in the 1:1 interviews, which lasted up to 1 hour. All of these were conducted virtually given many students had already left their student accommodations post-exams. These interviews were completely optional and offered to the full first cohort of student participants. Just as with the focus group, only certain questions from the 1:1 individual interview were used for this analysis. This included two questions that focused on social trust, and two questions that focused on institutional trust. This brings the total number of questions across the two data collection procedures to seven trust-related questions.

Although this limits the amount of text data that will undergo computational text analysis, it is still sufficient in size for the aim of this chapter. As discussed, this chapter is an attempt to show how computational text analysis may be applied to gender-based violence text data as an additional and more reproducible method compared to
qualitative work. This data-analysis triangulation helps decrease bias, strengthen confidence in my findings by using the findings of Chapter 4 as ground truth (Thurmond, 2011). As Grimmer & Stewart (2013, p. 275) suggest, the best way to validate a dictionary-based sentiment analysis method is through comparison of the human coded “gold standard”. Consequently, this chapter’s primary goal is a theoretical and methodological contribution, acting as a potential blueprint for future applications of computational text analysis in mixed-methods research. In the future, a larger text data corpus would be desired to amplify the empirical contribution. In this chapter though, I will discuss the findings in terms of social trust and institutional trust, grouping the questions from the focus groups and 1:1 interviews into these two themes as in Table 5.1.

Table 5.1: Focus Group and Individual 1:1 Interview Questions to be Analysed

<table>
<thead>
<tr>
<th>Interview Type</th>
<th>Social Trust Questions</th>
<th>Institutional Trust Questions</th>
</tr>
</thead>
</table>
| Focus Group Interview| • To what extent do you feel confident relying on fellow students to act as an active bystander if you find yourself in an uncomfortable situation?  
• What would you do if you saw someone [a fellow student] being sexually harassed? | • How do you think these [university] resources could be improved? |
| Individual 1:1 Interview | • Do you think the training should be mandatory for all students? Why or why not?  
• Do you think there is a common understanding of where that ‘line’ is amongst the student population? Why or why not? | • Do you feel the training is an adequate measure to combat gender-based violence at the university? Why or why not?  
• Do you think the training should be mandatory for all staff? Why or why not? |

As seen in Figure 5.2 and Figure 5.3, the sample is not overall representative of the university’s gender demographics. No males participated in the optional individual interviews and males in the focus groups were limited. Rather, 71% of the participants were female and the remainder of participants identified as either gender non-binary or transgender. As discussed in Chapter 3, the high proportion of female participants in both the focus groups and 1:1 individual interviews is expected as just about 60% of the university population identifies as female. While the sample is representative of the female population at the university, the sample is not for the male population nor transgender/non-binary population. This limits the generalisability of the findings as women are more likely to be attuned to gender-based violence given the higher rates of victimisation. Their trust in other students and the institution itself may not reflect all
students’ perspectives on social trust and institutional trust. Again, while this chapter makes an empirical contribution, the extent to which these findings may be considered widely applicable to the universal higher education student, regardless of gender, is limited.

In terms of other relevant demographics of the 1:1 interviews, the sample was majority non-heterosexual white students. Both the focus group sample and 1:1 interview sample had the exact same split in terms of race, with 71.43% of participants identifying as white. The remaining 28.57% identified as Asian. The focus group sample and 1:1 interview sample differed slightly in terms of students’ sexuality. The 1:1 interview participants were 14.29% heterosexual and 85.27% non-heterosexual. For the focus group sample used in this chapter, 21.43% of participants were heterosexual and 78.57% were non-heterosexual. This means that this chapter’s findings are likely skewed towards that of the non-heterosexual students’ perceptions of trust. A discussion of racial and ethnic representativeness of the focus groups may be found in Chapter 4. For a more detailed discussion of the overall racial and ethnic representativeness of the entire sample, please refer to Chapter 3.

**Figure 5.2: Gender Breakdown of Focus Groups**
Prior to discussing how analysis took place, it is pivotal to explain how the second text corpus was collected. I decided to use questionnaires as Alsaid et al. (2022) did to form the base for developing their trust lexicon. I did this because questionnaires can measure one’s perception of trust and different surveys present various ways that trust has been conceptualised (Alsaid et al., 2022). In collecting multiple questionnaires, I can comprehensively analyse how trust is conceptualised in a myriad of ways and as practitioners actually choose to measure it. While Alsaid et al. (2022) collected trust questionnaires focusing on automation, e-commerce, and interpersonal trust between humans, I adapted this of course to gender-based violence questionnaires. Additionally, rather than utilising whole questionnaires in my corpus as Alsaid et al. (2022) did, I collected survey items that focused on trust from 12 surveys. These surveys analysed trust as it relates to gender-based violence and came from a literature search done on Google Scholar using the keywords: “trust”, “sexual violence”, “gender-based violence”, “bystander sexual consent training”, “higher education”, and “university”.

While the search initially excluded any research not evaluating higher education gender-based violence, this was adjusted. The main reason for this is because the number of surveys explicitly developed to look at trust as it relates to gender-based violence are few in number. In fact, just two surveys found focused explicitly on trust and gender-based violence. These included Sulkowski’s (2011) Trust in College Support System Scale (TICSSS) and Holland et al.’s (2016) study of U.S. army member’s trust in the army system. Hence, both surveys were included in this study. A full list of the surveys included in the bespoke dictionary development can be seen in Table 5.4. It is important to note that some surveys pulled from other surveys’ items. For instance, “College officials handle incidents in a fair and responsible manner” was used in three different surveys. In this case, the duplicate items were kept capturing that words such
as “fair” and “responsible” should be weighted more heavily when developing the trust lexicon. The trust survey data contained 193 unique survey items.

Each survey was reviewed twice for questions relating to institutional trust and social trust, whether trust was explicitly mentioned or implicit in the question. An example of a survey item that explicitly addressed trust would be Holland et al.’s (2016) item “If you are sexually assaulted, you can trust the military system to protect your privacy”. An example of survey items exploring trust implicitly would be, for example, “My college does enough to protect the safety of students” or “There is a good support system on campus for students going through difficult times”. Both of these survey items came from the White House Task Force Climate Survey, as cited in Paulk et al. (2017). In total, 239 survey items were initially collected to be included in the computational text analysis. 27 of these survey items were later excluded as they were negatively phrased, leaving 212 positive survey items. An example of this would be Sulkowski’s (2011) survey item “The college responds too slowly in difficult situations”. While I debated reversing the negation of these survey items, I felt it would skew the words driving the lexicon development in a way that was inauthentic to how administrators are discussing these topics with students. As a result, I chose to remove these negatively phrased survey items, although in the future it may prove useful to create a lexicon with these survey items. In the end, 47 survey items related to social trust, while 165 survey items related to institutional trust.
<table>
<thead>
<tr>
<th>Survey Name (if applicable)</th>
<th>Source</th>
<th>Context</th>
<th>Number of Survey Items Included in Uni-Related Dictionary Development</th>
<th>Number of Survey Items Included in Student-Related Dictionary Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>N/A</td>
<td>Banyard et al. (2021)</td>
<td>Sexual violence on university campus</td>
<td>24</td>
<td>5</td>
</tr>
<tr>
<td>MIT Community Attitudes Survey</td>
<td>MIT (2014)</td>
<td>Sexual violence on university campus</td>
<td>24</td>
<td>10</td>
</tr>
<tr>
<td>Sexual Assault Campus Climate Survey</td>
<td>Higher Education Data Sorting (HEDS) (2023)</td>
<td>Sexual violence on university campus</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td>Rutgers iSpeak</td>
<td>Seabrook et al. (2018)</td>
<td>Sexual violence on university campus</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>Administrator Research Campus Climate Collaborative (ARC3)</td>
<td>ARC3 (2018)</td>
<td>Sexual violence on university campuses</td>
<td>24</td>
<td>4</td>
</tr>
<tr>
<td>UniSAFE</td>
<td>Lipinsky (2021)</td>
<td>Gender-based violence on university campuses</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Trust in the College Support System Scale (TICSSS)</td>
<td>Sulkowski (2011)</td>
<td>Violence on university campuses</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>N/A</td>
<td>Gonzalez (2022)</td>
<td>Sexual violence on university campuses</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Association of American Universities (AAU) Survey Questionnaire</td>
<td>Cantor et al. (2016)</td>
<td>Sexual violence on university campuses</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Department of Defense’s 2010 Workplace and Gender Relations Survey</td>
<td>Holland et al. (2016)</td>
<td>Gender-based violence in the army</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Bystander Behaviour Scale</td>
<td>Alegria &amp; Flores et al. (2017)</td>
<td>Gender-based violence on university campuses</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Bystander Confidence Scale</td>
<td></td>
<td>Gender-based violence on university campuses</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Willingness to Help Scale</td>
<td></td>
<td>Gender-based violence</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
5.2: Analysis

The intended outcomes for analysis are two-fold: to develop a bespoke dictionary and use that dictionary to run sentiment analysis on the data corpus from qualitative interviews. I would first like to explain what sentiment analysis is, and specifically how dictionary-based sentiment analysis is the best methodological choice. Accordingly, sentiment analysis is a type of computational text analysis used to evaluate the emotional charge of text data, in this case, conversational dialogue (Rice & Zorn, 2021). Some sentiment analysis evaluates positivity and negativity, while other sentiment analysis involves an evaluation of emotional states (ex. anger or trust) (Rice & Zorn, 2021). It assumes language can be used to indicate emotion without the word for the actual emotion being used (Rice & Zorn, 2021). The benefits of sentiment analysis is that it can offer insight without close reading. In the case of larger text corpora, sentiment analysis can be scaled with no increased amount of human effort.

There are two main options for sentiment analysis: machine learning powered analysis and dictionary-based analysis. Machine learning requires a large text corpus where a portion of the corpus is a training dataset (Rice & Zorn, 2021). Sometimes, human coding is brought into the training step in order to validate the classifier. Machine learning is unsuitable in the case of this research study given the size of my text corpus. On the other hand, dictionary-based sentiment analysis is completed by cross-checking words in the dictionary with words in the text corpus being evaluated (Benoit, 2020; Grimmer & Stewart, 2013). Pre-defined dictionaries, like AFINN, ANEW, and NRC, are well established pre-existing dictionaries. For example, ANEW is a dictionary made up of words that measure arousal, pleasure, and dominance. Yet, these dictionaries are developed to apply to general contexts, which sometimes results in the meaning being lost in specific contexts (Grimmer & Stewart, 2013). As Rice & Zorn (2021) write, the usage of the word ‘love’ in tennis is widely different from the general use of the word ‘love’. To overcome this contextual limitation, bespoke dictionaries are gaining traction as they involve developing a domain-specific dictionary to eliminate the misconstruing of words in highly specific contexts (Rice & Zorn, 2021). As I seek to evaluate trust, which only two known dictionaries evaluate, I have chosen to create a bespoke dictionary to evaluate trust within the niche context of gender-based violence in higher education institutions.

To develop a bespoke dictionary, I have chosen to utilise word embeddings, similar to Alsaid et al. (2022) and Rice & Zorn (2021). If I had a larger input text corpus, it would perhaps make sense to create a bespoke dictionary purely based off the most common words in the text corpus. However, again, given the size of my text corpus, I am choosing to incorporate word embeddings. Word embeddings are vector representations of co-occurring words within a text of collection of texts (Li et al.,
They are widely used in natural language processing (Li et al., 2020). Word embeddings may be used in two different ways, by applying them in-text, or leveraging pre-trained word embeddings. Within text word embeddings check the co-occurrence of words within a window of usually six words within the corpus itself (Li et al., 2020). It captures the nuance of language within the specific context accordingly. In contrast, pre-trained word embeddings look at global co-occurrences of words, which as shown below, offer a reliable way to expand a dictionary based off the most used or most unique words within a given text corpus (Alsaid et al., 2022). In developing their bespoke dictionaries, Rice & Zorn (2021) used in-text word embeddings, while Alsaid et al. (2022) used pre-trained word embeddings. I maintain a preference for pre-trained word embeddings in this instance for two reasons. First, questionnaire items, once cleaned, were sometimes too short for a six word in-text word embedding window. Secondly, I preferred utilising pre-trained word embeddings based on the top words from my text corpus to both consider the word’s context as well as the larger usage of the word in the English language (Pennington et al., 2014).

As this chapter is largely inspired by Alsaid et al.’s (2022) work, it is next necessary to review the steps in their computational text analysis while discussing my changes to their process. I will discuss these changes at a high-level and further detail on the more granular adaptations will be included later in this chapter. Alsaid et al.’s (2022) process involved first collecting questionnaires relevant to their inquiry, cleaning their text data corpus, and finding the top 20 most unique words by calculating the log odds ratio of those words in comparison to their use in the English language. From there, they calculated the cosine similarity distance for each of the 20 words to understand their similarity in semantic space to each other. They then collected the top five GloVe word embeddings, which will be explained below, for each of the 20 words too (Alsaid et al., 2022). I adapted their process by instead choosing the top words based on frequency in the text, rather than their uniqueness compared to the English language. I did this because, as mentioned, my surveys built on each other frequently. When survey items were repeated across different questionnaires, I wanted to capture that. I still decided to investigate the uniqueness of these words to validate that the words were uniquely looking at trust, but I did not utilise those ratings as determinants of whether the top words should be included in the dictionary or not. I also did not calculate the cosine similarity distance for each of the most used words. This is because I was less interested in how these words related in semantic space within the questionnaire text corpus, and sought to apply the lexicon to a different text corpus. For Alsaid et al. (2022), they were more focused on comparing questionnaire content, therefore cosine similarity distance made more sense for their research.

With these steps in mind, it was first necessary to clean the data from the 12 surveys. All steps of this and analysis were done with the R programming language (R
Development Core Team, 2016). The main R packages I used in pre-processing, cleaning, and analysis included: tidyverse (Wickham, 2016), tidytext (Silge and Robinson, 2016), ggplot2 (Wickham et al., 2019), textdata (Hvitfeld & Silge, 2021), and umap (Konopka, 2019). Unique to computational text analysis, pre-processing the text data corpus requires the destruction of the syntax of the questionnaire items (Benoit, 2020). To begin, the questionnaire items were converted to lowercase and 53 domain-specific words were removed. Removing domain-specific words was a step Alsaid et al. (2022) also took. Most of these domain-specific words were setting up the questionnaire items and included phrases such as “gender-based violence”, “police”, “institution”, “military”, “administrator”, “sexual violence”, “resident advisor”, etc. The plurals of phrases like these were also removed. Domain-specific words also taken out of the survey items were university names such as “MIT” and “Rutgers”. While domain-specific words are of course necessary to set up the survey questions, they are less critical to understanding how trust is conceptualised and demonstrated. The removal of domain-specific words was in addition to stop words and punctuation. Stop words are words that do not inform the text such as “the”, “of”, and “to” (Goldberg & Srivastava, 2022). With all of these steps, a survey item like “Policies at my institution which seek to tackle and eliminate gender-based violence are clear and explicit” instead reads “Policies at my which seek tackle and eliminate are clear and explicit”.

Next in the data pre-processing, I tokenised each survey item and then stemmed them. In this case, a token symbolised a one-unit word. Consequently, “Policies at my which seek tackle and eliminate are clear and explicit” was tokenised into 12 individual tokens. Stemming adds to the destruction of the syntax, but is a crucial step to identify patterns and pull meaning from the survey questionnaire items. Pre-stemming, the survey items consisted of 191 unique tokens. Stemming was done using the function stemDocument() in the tm R package. Critically, the function is limited in that, for instance, “discipline” and “disciplinary” would be stemmed into “disciplin” and “disciplinari”. For this reason, I manually stemmed the list of stemmed words further so that, in the example of “disciplin” and “disciplinari” both were stemmed to “disciplin”. This allowed the tokens to be counted as a group. Once final counts were taken, tokens were converted back to their original versions. To illustrate, “genuin” and “safeti” were rendered as “genuine” and “safety”. After the manual stemming, 168 unique tokens remained.

Next, I collected the top words from the list of unique tokens. As previously discussed, while Alsaid et al. (2022) used the top 20 most unique words as the basis for their bespoke dictionary, I decided to use the top 18 most frequently utilised words. The reason I chose 18 is because I wanted to ensure the bespoke dictionary remained small enough to be specialised to trust. These top 18 words each occurred at least seven times in the survey items. I manually inspected words beyond this cutoff to confirm that
this boundary made sense. Words that occurred six times did indeed feel less relevant to trust and included such tokens as “party”, “take”, and “channel”. I next calculated the log odds of the words compared to the wordfreq dataset. This compared how common the top 18 words were in the general English language (wordfreq dataset) compared to their frequency in the survey items. While this step has no impact on the bespoke dictionary development, it is still a crucial validation step. It provides a checkpoint to see how specific to the context of trust as it relates to gender-based violence the top 18 words are in comparison to their usage in everyday language.

I then used word embeddings to find the nearest neighbours for the top 18 words. Given the survey item dataset was small and each item was limited in terms of word count, I instead chose to use word embeddings in a pre-trained dataset. The Global Vectors for Word Representation (GloVe) dataset was built using Wikipedia pages and news text data from Gigaword. For each of the top 18 words in the survey items, I found the top 20 nearest neighbouring words from GloVe. Then, I chose 2-4 neighbouring words for each of the most common tokens to be included in the bespoke lexicon based on relevance to trust. The number of nearest neighbours I chose differed than in Alsaid et al.’s (2022) study as they chose the five nearest neighbours for each of their most unique words. I decided to incorporate more human decision making into this step based on Alsaid et al.’s (2022) own write-up of their limitations. With that, their final trust lexicon included nearest neighbours like “mathematics” and “phd” because of their automatic inclusion of the top five nearest neighbours from the GloVe dataset. Given I wanted to ensure my lexicon evaluated trust specifically, I decided to not automatically include the top five nearest neighbours. Rather, I read through the top ten nearest neighbours and chose 2-4. The final trust lexicon consisted of 65 words, including the original 18 most common tokens from the questionnaires. These were mapped into semantic space using Uniform Manifold Approximation and Projection (UMAP) which is a dimension reduction technique that uses, in this case, the 50 word embedding dimensions from the GloVe dataset. This step had no impact on which words were chosen in the final trust lexicon, but is an important validation step. By mapping each of the 65 words into semantic space, I could visually see how the words relate to each other. The UMAP therefore acts as a visual aid for understanding how trust is conceptualised.

I then conducted sentiment analysis using the bespoke trust dictionary. Sentiment analysis checks each word in the text data corpus being analysed to see if it matches tokens in the dictionary itself (Pennebaker et al., 2007). The total number of word matches is then divided by the total number of words in the text corpus to produce a percentage of the document considered as displaying sentiments of trust (Pennebaker et al., 2007). I compared the percentage results of the analysis to the percentage results from sentiment analysis using both Alsaid et al.’s (2022) trust
dictionary and the NRC dictionary’s trust lexicon. NRC is a pre-packaged dictionary consisting of 13,901 tokens which each correlate to either positive or negative emotion, or one of eight emotions. These emotions include anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. The trust tokens total 1,231 words, while Alsaid et al.’s (2022) lexicon has 100 tokens. This makes my bespoke trust lexicon the smallest trust lexicon of the three that I used for sentiment analysis.

It is important to note that I ran two individual sentiment analyses per dictionary. This is because I decided to split the text data corpus into two datasets. To do this, I combed through the questions asked in the 1:1 interviews and focus groups and pulled out the ones that were relevant to trust. Then, I categorised them as either targeting student-to-student (social) or student-to-institution (institutional) trust. The first dataset included 1:1 interview and focus group transcripts from questions targeting social trust. The second dataset included 1:1 interview and focus group transcripts from questions addressing institutional trust. The reason I did this is to ensure I could get specific trust estimates for both social trust and institutional trust, rather than one overall trust estimate. This allowed me to ground the sentiment analysis in my Chapter 4 findings. I knew from Chapter 4’s findings that social trust was higher than institutional trust. If a dictionary did not show this, I would know that the dictionary was not effective at estimating trust in this context of gender-based violence within higher education institutions. It is important to explicitly state that this chapter’s analysis was consequently conducted after both the data analysis and the writing of Chapter 4. This was a decision I made to ensure that this chapter’s findings did not influence my thematic analysis and survey analysis for Chapter 4, as well as to utilise Chapter 4’s findings as a ground truth for my sentiment analysis.

As mentioned, I removed domain-specific words prior to developing the bespoke dictionary. The bespoke trust lexicon ignores domain-specific words such as “institution” or “police”. The Alsaid et al. (2022) dictionary did this to some extent too, but the NRC dictionary contained domain-specific words. Hence, when I ran the sentiment analysis to cross-compare the dictionaries, it is notable that some domain-specific words found in the NRC dictionary boosted its overall sentiment score. This was less the case for the Alsaid et al. (2022) sentiment analysis and my own bespoke trust lexicon. Also different from the other two dictionaries, I ran my dictionary on the stemmed versions of the datasets given it was developed utilising stemmed words. The NRC and Alsaid et al. (2022) dictionaries were run on unstemmed versions of the datasets.
5.2.1: Top Stemmed Words

Figure 5.5: Top 18 Stemmed Words from Trust Questionnaires

As noted above, the first steps in analysis involved investigating word frequencies within the questionnaires and word frequencies of the top words compared to the English language. These steps were crucial in calibrating and validating the dictionary development by checking that these words were both relevant to trust and uniquely utilised at a greater frequency than in the general English language. As also discussed above, I decided to cut off the ‘top words’ at 18. The top 18 stemmed words were each stems that appeared at least seven times or more in the trust survey items. In Figure 5.5, it is clear that “report” is the most common stem at 47 occurrences. The next most common stems were “feel” and “safe” at 24 occurrences each. In general, all words seem to have to do with what would happen or what someone would do after a gender-based violence occurrence, except for the word “incident”. The majority of words could be considered verbs, including “feel”, “handle”, “inform”, “protect”, “respect”, “report”, “support”, “talk”, and “treat”. Three of these, namely “respect”, “report”, and “support” could also be nouns. This could show how trust as it is conceptualised in the questionnaire items, largely has to do with an action being taken. The adjectives associated with trust include “formal”, “fair”, “positive”, and “safe”. All of these do seem relevant to trust and how trust could be conceptualised in spoken word.

To compare the top 18 words’ uniqueness in the questionnaires to the English language, I took the log odds of each and compared them to the wordfreq dataset. Each of the 18 top stemmed words did in fact appear more frequently in the questionnaires than in the English language as seen in Figure 5.6. This is significant as in developing a
bespoke trust lexicon, it would be expected that the top words appear more frequently than in the English language. If one did not, it may be considered an additional stop word that could be removed as it is less likely to be relevant to trust specifically. It can be interpreted that the more unique of these 18 words, the more critical to identifying trust through the dictionary. “Report” which was the most counted word in the survey items, was actually the third most unique word in the survey items from the top 18 stemmed words. The most unique word in the survey, compared to the wordfreq dataset, was “talk” which only appeared nine times in the survey items. This was followed by “feel”, the second most common stemmed word in the questionnaires’ items.

**Figure 5.6: Uniqueness (Log Odds) of Top 18 Stemmed Words Compared to English Language**
5.2.2: Full Trust Lexicon from Word Embeddings

Figure 5.7: Trust Lexicon in Semantic Space Using UMAP

In Figure 5.7 above, the original 18 top stemmed words and their nearest neighbours are displayed through a UMAP dimension reduction. The dimension reduction is important because it offers insight into how the dictionary I developed conceptualises trust. The original 18 top stemmed words are symbolised by a black triangle, while the nearest neighbours collected from GloVe’s pre-trained word embeddings are represented by a black circle. First, the bottom to mid left of the scatterplot shows words that seemingly have to do with formal procedures. This includes words like “review”, “report”, “incident”, “official”, “press”, “response”, “steps”, and “agreement”. Most likely, students would not use these words when discussing their trust in other students. Rather, these words would likely come from conversations about institutional trust. The top left similarly has to do with more formal institutional aftercare including words like “treat”, “treatment”, “handle”, “protect”, and “help”. The top right of the scatterplot shows words that more likely have to do with social trust, such as “listen”, “tell”, “call”, “talk”, “feel”, “good”, and “happy”. Of course, these could still be related to institutional trust too, but lacks the formality and procedural tone of other words. Related to this, there are feel-good words in the middle right like “respect”, “fair”,
“dignity”, and “importance” that could emerge in both social trust and institutional trust conversations.

5.2.3: Sentiment Analysis

Figure 5.8: Full Word Cloud for Social Trust

Sentiment analysis of the interview responses first involved looking at word frequencies. Word frequencies offer an important and intermediary step prior to calculating the overall sentiment percentages. This step thus acts as a checkpoint because word frequencies can show how different dictionaries conceptualise trust through their unique lexicons. Word clouds are a common visual aid in computational text analysis for understanding word frequencies. The differences between the dictionaries are clear in looking at the word cloud in Figure 5.8. When analysing the text with the NRC dictionary, I detected 79 total tokens from the NRC dictionary and 100 total tokens from the Alsaid et al. (2022) dictionary. However, there were only 35 unique words for NRC and 26 unique tokens for Alsaid et al. (2022). The bespoke trust lexicon found 99 total trust words in the data corpus, but there were only 19 unique words. This difference is meaningful because it highlights how Alsaid et al. (2022) and the bespoke trust lexicon may be more tailored to analysing trust as it relates to gender-based violence given the sheer number of words found. Yet, NRC’s higher number of unique words reflects the fact that it maintains an extensive trust lexicon with 1,231 trust words. The bespoke trust lexicon has a mere 65 and Alsaid et al. (2022) has 120. It is worth a reminder here that the NRC dictionary and Alsaid et al. (2022) dictionary do not run on stem words. This is a major difference between the bespoke trust lexicon developed and the other two lexicons. While “confident” and “confidence” would be stemmed both to “confident” in the bespoke lexicon, the NRC dictionary maintains them
as two unique word tokens. The same can be seen with “trust” and “trusted” in the Alsaid et al. (2022) dictionary. This is important to keep in mind when comparing the number of unique words found by both sets of dictionaries.

A few noteworthy observations emerge when comparing the dictionaries. First, it is easy to see from the word clouds above that the Alsaid et al. (2022) dictionary and the bespoke trust lexicon both heavily rely on the term “feel”. This may be explained by the fact that both dictionaries were developed using questionnaires, where “feel” may be used to set-up the phrases in Likert scale questionnaire items. In fact, of the 99 bespoke trust lexicon words found, 38 of them were the word “feel”. The MIT survey, for example, had 11 survey items where the statement read, “I felt ______”. This was followed by “strong”, “confident”, “like I was getting closure”, “supported”, “safe”, etc. Interestingly, the majority of these words do not appear in the bespoke trust lexicon nor Alsaid et al. (2022) word clouds. Rather, the NRC dictionary found more of these trust-related adjectives. These included not just “safe”, which the trust lexicon and Alsaid et al. (2022) include too, but also “confident”, “personal”, “friendly”, and “understanding”. Additionally, the NRC dictionary found domain-specific words, while one of the first steps in developing the bespoke and Alsaid et al. (2022) lexicons was removing domain-specific words. These domain-specific words in the NRC dictionary include “school”, “officer”, and “handbook”. Two of these, “school” and “officer” were removed in the initial data cleaning of the trust survey questionnaires. I argue that “school” and “officer” are not universally trustworthy entities for everyone, especially students who may identify as part of a historically marginalised group. To illustrate, it is much easier to agree that “safe”, found across all three sets of dictionaries, is universally associated with trust. Another similarity in the words found included “depend” in the NRC dictionary, which can be likened to “rely” in the trust lexicon and Alsaid et al. (2022) lexicon. All of this goes to show that the dictionaries conceptualise trust in similar ways when it comes to social trust in particular, but with some key differences that may affect the final sentiment analysis scores.
For institutional trust seen in Figure 5.9, the NRC dictionary found the highest number of unique words again, and the Alsaid et al. (2022) dictionary found the least number of total words. With that, the NRC dictionary found 50 total words, the bespoke trust lexicon found 51, and Alsaid et al.’s (2022) dictionary found 27 total tokens. The total counts across the dictionaries are lower across all three dictionaries which potentially hints at lower levels of institutional trust than social trust, but this needs to be weighted by total word counts. In terms of unique words, NRC did find just about double of them than the other two dictionaries. NRC found 31 unique words, while my bespoke dictionary found 15 unique words and Alsaid et al. (2022) found just 13. Altogether, this shows how NRC’s more extensive dictionary may lead to higher counts of unique words. Yet, lower total counts and unique word counts across all three dictionaries show how institutional trust may be lower than social trust across the board. It is also likely that Alsaid et al. (2022)’s dictionary may show the lowest levels of institutional trust given it picked up on just 27 total tokens compared to NRC’s 51 total tokens and the bespoke dictionary’s 50 total tokens.

A few patterns also emerge in the word-level analysis of institutional trust. To demonstrate this, the NRC once again found more domain-specific words. This included “school” as it did in the previous analysis, but also “police”, “authority”, and “centre”. I once again argue that the police are not a clearly trustworthy space for many people from historically marginalised groups. This was echoed by a participant in particular who stated that “[the university administrators] do forget that for a lot of people, the police are not a safe place.” Related to this, another substantial difference in the conceptualisation of trust at a word-level is that the bespoke lexicon is more adjective and verb focused than the other two dictionaries. NRC’s word clouds have action words such as “providing”, “manage”, “deal” and “believed”. Alsaid et al.’s (2022) lexicon has “manage” and “providing” too. The bespoke dictionary has the same or similar action...
words such as “provide”, “handle”, and “respond”. This is in addition to other verbs too though such as “talk”, “report”, “consult”, and “support”. Notably, all three dictionaries have some version of “provide” and “safe”, showing that this is likely important when conceptualising trust. Furthermore, “resource” appears in the bespoke and NRC word clouds, while “information” is in the Alsaid et al. (2022) word cloud. This highlights the interconnectedness between trust and sources of help, whether resources or basic information.

**Figure 5.10: Trust Sentiment of Interviews (%) per Dictionary**

![Figure 5.10](image)

Figure 5.10 shows the sentiment scores as percentages for both the social trust and institutional trust datasets once the dictionary token counts were weighted by the total words in each dataset. Given I learned social trust is higher than institutional trust in my thematic and survey analysis in Chapter 4, I use this finding as my ground truth in evaluating the dictionary sentiment analysis. In this way, I can calibrate the findings in Figure 5.10 to see that the NRC trust dictionary is the only dictionary to incorrectly assign magnitudes to social trust and institutional trust. Both non-off-the-shelf dictionaries correctly estimate that social trust is higher than institutional trust. NRC either underestimates social trust or overestimates institutional trust. Therefore, while the NRC dictionary is a frontrunning dictionary in computational text analysis, in this case, it is worthwhile to utilise more bespoke dictionaries. This could include developing a bespoke dictionary such as the bespoke trust lexicon or incorporating existing...
bespoke lexicons such as Alsaid et al.’s (2022). Out of curiosity, I also conducted sentiment analysis for the bespoke trust lexicon on both datasets, but without stemming the interview data. This still output social trust higher than institutional trust, as it was found to be in Chapter 4.

There are a few reasons for why the bespoke trust lexicons could be more accurately estimating trust in both the interview datasets. Primarily, there is a major difference in how the NRC dictionary was trained and how the bespoke dictionaries were developed. NRC was built through Mechanical Turk whereby humans completed word association tasks to determine if and how words relate to trust. This was the largest and first dictionary of its kind. The two other dictionaries were trained using questionnaire trust items. Alsaid et al.’s (2022) lexicon was developed with 626 trust questionnaire items looking not specifically at how it relates to gender-based violence, while my bespoke trust lexicon did just that with 165 trust questionnaire items. Both offer different, but comparably reliable approaches to building a trust dictionary for sentiment analysis.

The fact that bespoke dictionaries are more reliable and reproducible in this case prompts the question of whether smaller dictionaries are more worthwhile to develop. The size of the two dictionaries were significantly smaller than NRC’s trust dictionary, with NRC maintaining 1,231 trust-related tokens. The potential benefit of my bespoke dictionary’s mere 65 tokens, and Alsaid et al.’s (2022) 120 means that the analysis could be more specific to the context. As mentioned, while Alsaid et al.’s (2022) lexicon was developed for ecommerce and automation, it still did also focus on human-to-human contexts too. My dictionary took this a step further by specialising even more on gender-based violence, yet producing similar results that reflected the findings of Chapter 4 too. This shows how perhaps it is not necessary to invest as much human effort as in NRC’s dictionary development. Rather, this chapter shows how there is enough reason to think about how we can leverage text documents to aid in smaller dictionary development.

5.3: Discussion

The main methodological finding that in itself is a contribution to the literature is that computational text analysis can be applied to gender-based violence research to measure trust. This is important as future evaluations of training interventions can incorporate computational text analysis to measure pre- and post-implementation changes. This chapter offers a method for analysis that could be expanded to do just that. Rather than trying to glean students’ perceptions solely through thematic and survey analysis, I offer an additional and reproducible method that refocuses on participants’ behaviours and what was specifically said. While Alsaid et al. (2022) set
the pathway for coming up with sentiment-specific dictionaries based on trust questionnaires, this chapter takes their work one step further by applying an adapted dictionary development process to actual sentiment analysis of real-life interview text data. I therefore adopt their dictionary development process by making it more domain-specific to the gender-based violence context. Even though this process results in a smaller dictionary, this work shows how a reliable dictionary can still be developed. This remains true even as domain-specific surveys can be limited, as trust surveys in the gender-based violence space are. Consequently, this chapter demonstrates how bespoke sentiment lexicons can be developed regardless of field, as long as there are enough relevant training documents. These can be questionnaires as they were here and in Alsaid et al. (2022), or other text training data such as websites, speeches, online forums, etc. This chapter contributes to the possibility that more computational text analysis is worthwhile in the field of gender-based violence and beyond.

Related to this, the comparison of the bespoke trust lexicon to Alsaid et al.’s (2022) trust dictionary and the NRC trust dictionary was critical in this chapter. The NRC dictionary is a go-to dictionary for researchers seeking to run sentiment analysis. However, this bespoke trust lexicon showed how other dictionaries, just waiting to be developed, may be more accurate at measuring sentiment in text. I thus demonstrate the reliability of my bespoke trust lexicon by grounding my findings in my previous work. More dictionaries do not equate to the loss in the use of the NRC dictionary. Rather, multiple dictionaries can be used and compared to fully ensure the levels of sentiment are accurate. Together, the dictionaries are powerful tools to investigate more covert group-level sentiments as they relate to gender-based violence. In grounding the reliability of the results of computational text analysis in my previous work, I exhibit how dictionaries can be a standalone methodology or utilised as an additional layer of analysis. The latter may be used to reinforce qualitative findings, such as thematic analysis, and render the findings of inequality regimes more reproducible.

I made two key changes to Alsaid et al.’s (2022) process to create a more domain-specific lexicon. The first, is that I decided to limit the questionnaires I collected to gender-based violence questionnaires. Given there are not enough trust-specific gender-based violence questionnaires (just two to my knowledge), I expanded my search for GBV questionnaires to any evaluating trust, explicitly or implicitly, in at least one questionnaire item. This made my lexicon more domain-specific from the first step of development. My second adaptation to Alsaid et al.’s (2022) process was utilising the training words for my dictionary from the top words based on frequency, rather than their uniqueness compared to their use in the English language. I believe this was crucial in the development of my dictionary as it ensured I considered that the limited number of trust-related items in gender-based violence questionnaires often built on each other from one questionnaire to another. Overall, my process was guided by my
desire to create a truly domain-specific dictionary that was less concerned with the words’ relation to each other in semantic space within the training documents. I was less concerned with how the questionnaires differed from each other given they were all part of the same domain. The values of these changes could help explain the dictionary’s ability to accurately estimate social trust as higher than institutional trust in the final sentiment analysis, even given the limited number of training documents.

While the computational text analysis method I developed here is valuable in reproducing the findings of Chapter 4, there are also evident options for refining the dictionary development and utilisation. For instance, one idea for refinement could include testing the removal of words which I refer to as “set-up words”. To me, set-up words are words used to create the foundation for the more important words in the questionnaire items. As an example, “feel” is one that appears frequently in both sentiment analyses for my own bespoke trust lexicon and Alsaid et al.’s (2022) lexicon, as seen in the word clouds in Figure 5.8 and Figure 5.9. The removal of the word could be explored in the step where domain-specific and stop words were removed. To illustrate this, the MIT survey used “feel” or “felt” as a set-up word 13 times, such as in “I felt strong” or “I felt supported”. Other set-up words like “happen” or “incident” could also be potentially removed to see how overall sentiment scores are affected. Another potential step for refinement could have to do with stemming words. While it is beneficial to stem words in the questionnaire survey item steps to understand what the top words are, it may be less necessary to keep them stemmed throughout the dictionary development. Instead, the dictionary could include all versions of a word. An example could be “support”, in addition to “supported”, “supporting”, “supportive”, “supports”, “supporter”, etc. This is what NRC does and perhaps it does add some nuance to the sentiment analysis in that tenses can be telling for whether students feel (dis)trust currently or in the past.

Other potential refinement steps could include adding the word “trust” itself, or even adding in other trust dictionaries into the training documents. It is important to say that it was not a conscious decision to not add the word “trust” itself into my bespoke trust lexicon. “Trust” simply was not in the top 18 stemmed words from the questionnaire items. Rather, it was #33 out of #256 stemmed tokens. It had just five appearances, while the cut off for the top 18 stemmed words was seven appearances or more. I do feel that there is some logic to not including the word trust. After all, this research seeks to find covert sentiments of trust as it relates to gender-based violence. By including the word “trust”, this may more explicitly uncover overt sentiment rather than the covert. Even so, if “trust” was to be added, another option could be including the NRC lexicon or Alsaid et al.’s (2022) trust dictionary into the training documents. This may help create a more comprehensive bespoke trust lexicon while still grounding the majority of the dictionary in the gender-based violence questionnaires.
In the future, it would be worthwhile to expand the datasets used for this computational text analysis to further validate the development process. I want to acknowledge that the datasets I used sentiment analysis on were small, while the participant samples were not perfectly representative either. As mentioned, students participating voluntarily in the student are likely to be a select group of individuals. I also used text analysis on responses from just three questions for institutional trust and just four questions for social trust. Furthermore, I did not do any type of pre- and post-analysis with my computational text analysis because the text corpora would be even smaller. For that reason, it would be worthwhile to measure trust in the same way in the future, but with more expansive text corpora. Embedding a text corpus where levels of trust are more known could help create a threshold for (dis)trust when it comes to sentiment analysis as well. Expanding potential corpora beyond focus groups could include forums where students are discussing gender-based violence online, or even transcripts of victim-survivors’ statements to the university’s disciplinary team. The latter may have higher levels of distrust, which could be utilised to benchmark sentiment analysis of the other text corpora. Similarly, it may prove methodologically critical to ask the same questions for evaluating social trust as institutional trust in the future too. In this chapter, I compare social trust and institutional trust, but their text corpora were produced from different questions. The precise percentages found likely maintain some error given this, but it remains worthwhile that the two bespoke dictionaries accurately estimated that social trust was higher than institutional trust.

The main empirical and theoretical contributions of this chapter echo that of Chapter 4. This chapter once again reveals how it is worthwhile to investigate group-level covert manifestations of trust as they relate to gender-based violence. While surveys offer pulse checks of how the student body is feeling, computational text analysis is a quick way to analyse these more covert sentiments that have real impact on students’ actions. It also further shows that if universities wish to truly mitigate gender-based violence, it must be investigating the differences between social trust and institutional trust. Although there may be some error in the exact percentages found with the sentiment analysis, the fact that all three dictionaries showed that trust sits at just about 10% and less is alarming. This in and of itself should act as a startling wakeup call for the university. To me, these low percentages highlight how the training intervention cannot be a one and done solution.

Moving forward, the findings from Chapter 4 echoed here mean that universities must reckon with the low levels of trust students have in the institution and the highly contextual trust they have in other students. Again, the training itself is a positive value-add, but its value-add is limited. It boosts students’ confidence in themselves and others’ intervention tactics, but an overall trustworthy culture still lacks. Chapter 4 details what universities could be doing to expand the trust students have in the
institution, as well as in the general student body. These included expanding student-to-
student trust building through social programming, as well as working with third party
organisations that are not under the university’s umbrella. One small initiative that was
started in one school of the university is including a link to the gender-based violence
reporting and supportive resources platform in student-facing classroom documents.
For instance, tutors have been equipped with a resource slide to include in any
presentation they give to students. Granted, that is just one school targeting one portion
of the larger student population. Another small initiative in the works seems to be the
idea of booster sessions. While the conversations around booster sessions are just
that, conversations, these could be a worthwhile initiative. Such sessions offered to
students halfway through university could offer a reminder of the trust students could
have in each other and in the institution. This also may prove worthwhile as the training
intervention is mostly targeted at first years who are likely overburdened with other
training programming.

The university at this point now knows there is a (dis)trust students have in the
institution preventing and effectively managing gender-based violence. The exact
reasons for why were not within the scope of this research study, but need to be further
explored. This could include additional 1:1 interviews and focus groups with students
who are not just first year students. This is critical as this study is reporting on (dis)trust
of first year students, which means the lack of trust could be even worse for students
who have more university experience. The university could try to recruit not just
students passionate or previously unknowledgeable about the topic, but should also
place an emphasis on victim-survivors if they are willing to participate. Questions could
include specifically asking students “What would help you, as a student, trust other
students more when it comes to gender-based violence?” and “What would help you, as
a student, trust the university more when it comes to managing and preventing gender-
based violence incidents?”. Crucially, further research could include the university
directly asking students for recommendations of organisations they respect who the
university should look to for best practices and other training interventions.

It is this lack of institution to student dialogue that concerns the pessimist in me.
While it is obvious to me that both further action and further research are warranted, I
do believe students need to be at the heart of any next steps. If institutional trust is so
low, as Chapter 4 finds and this chapter reinforces, why is the university going to the
industry market or solely speaking to other universities instead of going directly to the
source to understand how to boost student trust? Prior & de Heer (2021) write of
universities being plagued by bureaucracy, an overwhelming concern for brand
reputation, and finances, rather than the wellbeing of the end consumer: the students. I
write in Chapter 3 of the limited incentives for students to get involved in my research,
which ultimately has contributed to my small data corpora here. Without an investment
in dialogue with as many students as possible and a sense of urgency (not just stemming from administrators feeling pressure from active student campaign groups), I am fearful any strategies to boost student trust as it relates to gender-based violence will struggle.

5.4: Limitations

Some limitations involve the methodology itself. Mohammadi & Hashemi Golpayegani (2021, p. 2033) write, “Trust is context-specific, meaning that it is always bounded to a particular context”. This quote highlights the limitation that trust was not directly asked about in the focus groups or one-on-one interviews. Rather, computational text analysis was used to uncover covert sentiments of trust at the group-level. If this study were repeated, it may be worthwhile to incorporate interview data that comes from interview questions specifically naming trust, in addition to using the same questions for both social trust and institutional trust. This chapter has proven that there is a benefit to this methodology, but I will acknowledge that speaking to students directly as suggested above, is an advantageous next step. This will only help get conversations started in the university community, and only then can the problem of gender-based violence in the university context be effectively addressed and mitigated. Likewise, more male students should be recruited in future research. This study had a majority female participant population, which is reflective of the university itself. However, for the 1:1 interviews, there was a complete absence of male participants. To specifically target potential male participants, the university could reach out to male-only sports clubs or social clubs. Even beyond male participants, the research could benefit from more participants in general, which in turn could be broken into pre- and post-text data corpora.

In terms of the lexicon development, a few other limitations emerge. First, I've noted that the number of trust surveys relating to gender-based violence were inherently limited. As a result, I chose to select what I deemed to be relevant survey items relating to trust. This could be seen to introduce some bias in the lexicon development process as others may find the survey items I chose may not have to do directly or indirectly with trust. Another solution could have been casting a wider net for trust surveys, rather than trying to stick to gender-based violence surveys. Other surveys could measure trust within neighbourhoods, in the workplace, in online communities, in politics, etc. Limitations may also include the decision to remove domain words and to change the cutoff of the top stemmed words. The top stemmed words from the survey items were limited to 18 words with seven appearances in the survey items each. These decisions were for two reasons. First, Alsaid et al. (2022) stuck to 20 words and similarly removed domain-specific words. Secondly, I inspected the list of top stemmed words and words
with six appearances did in fact seem less relevant to the concept of trust. They included words such as “party”, “take”, and “leave”. I stuck to the top 18 stemmed words which seemed significantly more related to the concept of trust. For domain words, the words I chose could have been confirmed to be irrelevant through some kind of Mechanical Turk association task. This could have been similarly done for the words I chose from the word embeddings’ nearest neighbour lists too. I chose the top 2-4 words I found to be most relevant to the original word from the nearest neighbours list, but this is of course susceptible to bias too. Given I had no funding for this research, I stuck to my best judgement. I feel the decisions I made throughout the computational text analysis process proved to be small limitations in the grand scheme. This is because, in the end, the sentiment analysis did reproduce similar conclusions to that of Chapter 4.

5.5: Chapter Conclusion

This chapter explores covert sentiments of trust at the group-level in a different, but equally valuable way as to Chapter 4. A bespoke trust lexicon was developed using computational text analysis through word embeddings and sentiment analysis. Utilising the new dictionary in conjunction with the NRC dictionary and Alsaid et al. (2022) lexicon shows how a bespoke lexicon is worthwhile. Simply, although sentiment dictionaries already exist and are used in the computational social science community does not mean they should not be challenged. Furthermore, this methodological contribution offers up the opportunity for further refinement of the dictionary with adaptations made to Alsaid et al.’s dictionary development process. Such refinement could include the incorporation of additional training documents as well as the application of the dictionary development process to other fields beyond that of gender-based violence. Empirically, this chapter finds once again that social trust is higher than institutional trust. Suggestions for universities to move the needle in terms of trust could include such programming as extra-university partnerships, booster training interventions, and most critically, further research. In the end, this chapter emphasises that the problem of student (dis)trust cannot be solved without the continued involvement of students. Hearing their thoughts, opinions, stories, and voices must continue and be placed as a higher priority than brand management and bureaucratic procedures. This must be the first step, and undoubtedly would also create more text data that could be analysed. It is my hope that further research will be conducted to address these (dis)trust issues in higher education, all the while researchers build computational text analysis into their analyses.
Chapter 6: Quantifying Inclusion with Survey Analysis

Keeping in mind the existing inequality regimes within academic spaces, this chapter arises from the desire to investigate how patterns of inclusion/exclusion shift over time within the university’s classrooms. It takes a survey-based approach to inclusion and exclusion by researching student perceptions of these phenomena to understand if any one demographic group in particular feels more excluded than others at two timepoints after the administration of the inclusion training interventions. The demographic groups directly studied in this analysis are based on gender, race and ethnicity, disability, sexuality, involvement in university extracurriculars, school, and school year, as each are hypothesised to potentially impact a student’s experience of inclusion. While not a comprehensive study of all dimensions of diversity that could lead to covert and overt exclusion of groups of students in the classroom, this chapter still offers an empirical and methodological contribution to the literature.

With that, this chapter operationalises inclusion with the Mor Barak Inclusion-Exclusion (MBIE) scale by adapting it for students in two different university schools. This study is the first of its kind to adapt the MBIE to the education space. Of course, blatant discrimination of whole demographic groups by tutors is uncommon, but the MBIE scale allows for an investigation into how perceptions of inclusion, or exclusion, are aggregated at the group-level. As discussed in Chapter 2, the literature generally over-relied on perceptions of inclusion, but understanding these group-level perceptions of inclusion are still necessary to provide practical feedback for finetuning of the interventions previously administered. Furthermore, this chapter’s focus on attitudinal inclusion seeks to offer a complementary analysis to Chapter 7’s study of observable behaviours of inclusion/exclusion change at the same timepoints within the same classroom context. Together, they exemplify my new theoretical framework that seeks to minimise the overreliance on perceptions of inclusion/exclusion by inquiring into both perceptions and behaviours. This chapter is therefore an important first step in establishing how covert patterns of inclusion/exclusion may exist at the group-level.

To achieve this, I use statistical analysis of the adapted MBIE survey. Using statistical analysis allowed overall sentiments of inclusion/exclusion per demographic group to be explored. Empirically, this chapter failed to find evidence of a blatant hierarchy or inequality regime within the classrooms for the demographic groups studied. A key finding though is that perceptions of inclusion were significantly higher later in the term than earlier in the term. Even so, I do reveal how race and ethnicity are the most salient dimensions of diversity when it comes to significant differences in experiences of inclusion in the classroom. This adds to the literature by showing how racial and ethnic minorities may experience small group educational settings differently. While a direct link to the inclusion trainings administered cannot be drawn
given sample and research design limitations, item-by-item analysis also shows how the trainings may be augmented to better foster inclusion for all demographic groups.

In concluding this chapter, I strive to write this chapter as a feedback loop. To date, there is no agreed upon meaning of what inclusive teaching practices are (Finkelstein et al., 2021). With this survey administration and subsequent statistical analysis, it is possible to point to the exact behaviours that made students feel less included. By bringing in anecdotal tutor perspectives, I highlight how there may be a disconnect between what tutors think they are doing and what they should be doing to foster inclusion. In relation to Chapter 7, I discuss how these findings may be situated within the literature that exists on international students as a significant portion of the Asian student participants discussed in this chapter identified as Chinese students.

Beyond race and ethnicity, I emphasise how important it is for universities to incorporate inclusion surveys, such as the adapted MBIE scale survey, to continue to understand how all students are perceiving inclusion at the group-level. In turn, this information can affect the development and reinforcement of inclusion trainings for tutors and all teaching staff.

6.1: Data

Data were collected from spring 2022 through autumn 2022. 249 students and seven tutors participated in this portion of the study. Students and tutors came from two different schools with the inclusion training tutors received dependent on their school. 92 students and three tutors came from School A, and 157 students and four tutors came from School B. Thus, a grand total of 256 participants were involved. The seven tutors had a total of 27 tutorial groups, which means the students were split into 10 tutorial groups in School A and 17 tutorial groups in School B. However, not every student in each tutorial group consented to participating in the study. An average number of nine students per tutorial group consented to participating meaning they took the survey twice during the term, while those who did not consent never received the survey link. Study participation did not translate into tutorial attendance with 60 total absences during the week of Survey #1, and 79 absences during the week of Survey #2. Attendance in both schools had no impact on students' final course marks.

As discussed in Chapter 3's methodology, the two surveys were administered through Qualtrics during Week #5 and either Week #10 or #11 (dependent on the school). The time between surveys was an attempt to understand sentiments of inclusion towards the beginning of the term and the end of the term. At least two follow-up emails were sent per tutorial group to remind students to fill out each survey. Surveys remained open for exactly one week. Any responses submitted after the scheduled tutorial time one week after the survey opened were disregarded. This was
important to ensure experiences in the following week’s tutorial did not impact the responses about how students were feeling during the data collection weeks, especially for Survey #1. For Survey #2, this would only be the case for tutorial groups where tutors offered additional tutorial meetings to discuss the final essays or presentations. In total, seven late responses for Survey #1 were disregarded, while nine were excluded for Survey #2 for the same reason. Students used their student ID numbers to submit their survey responses. The responses were coded to each participant’s unique participant ID number. This combined their tutorial letter code (ex. A) with a randomised number to ensure anonymity in analysis.

Surveys took less than 15 minutes to complete all three sections. The first section asked about demographics. The second section asked about perceptions of safety, trust, and community (adapted from Unger & Wandersman et al., 1982). While these scales are not the focus of this chapter, I bring the responses in where relevant contextually. It is important to note too that these scales are on a Likert scale of 1 to 5 (1= Completely disagree, 5= Completely agree). Finally, the third section and the basis of this chapter’s analysis, focused on an adapted Mor Barak Inclusion-Exclusion Scale Likert survey (Mor Barak and Cherin, 1998). The original MBIE assessment consisted of 15 survey items with a 6-point Likert scale and was developed to measure perceptions of contribution and participation in work group processes. This version of the MBIE survey had only 11 questions for students and seven questions for tutors. Survey items dropped included, for instance, “I have influence in decisions taken by my work group regarding our tasks”. Items like this could not be adapted to fit the context of tutorial groups as students generally do not have any influence in their tutorial agendas. That is solely left up to the course organisers and sometimes tutors themselves. In contrast, survey items like “I frequently receive communication from management higher than my immediate supervisor” could be adapted to “I frequently receive communication from my tutor”. The adapted MBIE scale was used to measure sentiments of inclusion/exclusion, as they relate to perceptions of contribution and participation.

The version of the survey for tutors similarly only consisted of questions that focused on their perception of inclusion in the tutorial group itself, not in the larger teaching staff group. For a full list of survey items used in the student and tutor surveys, please see Appendix C.3. Tutors were asked to fill out the survey keeping in mind the tutorial group where they had the largest number of students consent to participate in the study. This is because tutors sometimes taught up to five tutorial groups for the same course, but may have had as few as one student consent to being a participant. Tutors’ responses to the surveys will only be used anecdotally in this chapter due to their small sample size (n= 7). I will use these responses to aid in my findings of student perceptions of inclusion. Both students and tutors answered the questions on a 6-point Likert scale (1= Completely disagree, 2= Moderately disagree, 3= Slightly disagree,
4 = Slightly agree, 5 = Moderately agree, 6 = Completely agree). With no neutral response or non-response (N/A) option, the scale attempted to encourage students to critically think about their responses. The survey items for the student surveys may be seen in Table 6.1. The table provides an overview of the statements themselves, who the statement involves (i.e. tutor, peer, or unspecified), and the reference tag I will use to refer to these statements throughout analysis.

**Table 6.1: Overview of MBIE Likert Scale Statements**

<table>
<thead>
<tr>
<th>Survey Item #</th>
<th>Reference</th>
<th>Parties Statement Involves</th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Share Information</td>
<td>Peers</td>
<td>My peers in my tutorial group openly share information with me.</td>
</tr>
<tr>
<td>2</td>
<td>Invited to Tutorial Activities</td>
<td>Peers</td>
<td>I am typically involved and invited to actively participate in activities in my tutorial group by my peers.</td>
</tr>
<tr>
<td>3</td>
<td>Informed about Informal Activities</td>
<td>Peers</td>
<td>I am always informed about informal social activities and events with my peers in my tutorial group.</td>
</tr>
<tr>
<td>4</td>
<td>Invited to Informal Activities</td>
<td>Peers</td>
<td>I am often invited to join my peers from my tutorial group when they go out for a meal, drinks, or a study session.</td>
</tr>
<tr>
<td>5</td>
<td>Contribute Opinion</td>
<td>Tutor</td>
<td>I am often invited to contribute my opinion in tutorials by my tutor.</td>
</tr>
<tr>
<td>6</td>
<td>Last to Know</td>
<td>Unspecified</td>
<td>I am usually among the last to know about important changes in the tutorial.</td>
</tr>
<tr>
<td>7</td>
<td>Opinion Before Decisions</td>
<td>Tutor</td>
<td>My tutor often asks for my opinion before making important decisions.</td>
</tr>
<tr>
<td>8</td>
<td>Receive Frequent Communication</td>
<td>Tutor</td>
<td>I frequently receive communication from my tutor.</td>
</tr>
<tr>
<td>9</td>
<td>Participate in Communication</td>
<td>Tutor</td>
<td>I am invited to actively participate in communication with my tutor.</td>
</tr>
<tr>
<td>10</td>
<td>Influence Decisions</td>
<td>Unspecified</td>
<td>I am able to influence decisions that affect my tutorial.</td>
</tr>
<tr>
<td>11</td>
<td>Outside Meeting Invites</td>
<td>Tutor</td>
<td>When there is a meeting with the course organisers outside of my tutorial, I am invited to participate by my tutor.</td>
</tr>
</tbody>
</table>

**Table 6.2: Total Student Survey Completions per School**

<table>
<thead>
<tr>
<th>School</th>
<th>Number of Students</th>
<th>Survey #1</th>
<th>Survey #2 (Δ = Difference from Survey #1 to Survey #2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>92</td>
<td>55.43%</td>
<td>38.04% (Δ = -17.39%)</td>
</tr>
<tr>
<td>B</td>
<td>157</td>
<td>63.69%</td>
<td>51.59% (Δ = -12.10%)</td>
</tr>
</tbody>
</table>
Of the 27 tutorial groups participating in the study, only 11 tutorial groups maintained or improved their completion rates from Survey #1 to Survey #2. The average completion for Survey #1 was 61.40% and dropped to 49.48% for Survey #2 for students. For a full overview of each tutorial group’s survey completion rates, please see Appendix C.4. Tutors had 100% survey completion rates for both surveys. This is likely because tutors personally knew me and knew the study was for my PhD research. So, when reminder emails came from me, they perhaps felt more inclined to respond than students who had never met me in-person. As seen in Table 6.2, both courses saw drops in student completion rates, with the drop for School A being slightly higher than the drop witnessed for School B’s student survey completion rate. The lower completion rate for School A is perhaps for a few reasons. Survey #2 was sent to School A in Week #11, instead of Week #10. This is because the school had a mid-term break where students had no tutorials. Survey #2 was therefore sent a full week closer to the end of term and closer to the final exam period. Anecdotally, a tutor also suggested that attendance may have dropped due to FIFA World Cup matches coinciding with tutorial schedules. Hence, due to conflicting interests and exam pressure, students may have failed to fill out Survey #2 in both schools, but especially in School A.

Another potential reason for a drop in survey completion rates may have been due to the University and College Union (UCU) strikes. These strikes were conducted to advocate against universities’ pay and working conditions. Tutors who are UCU members or who went on strike in solidarity with union members withheld tutorial sessions on strike days. Of the 27 tutorial groups involved in this study, nine tutorial groups’ sessions fell on nation-wide strike days. 51 students from these nine tutorial groups filled out Survey #2. Only seven students reported they still attended their regularly scheduled tutorial session, and 44 students reported they did not attend their tutorial nor any other session in lieu of their own. I used a Welch two-sample T-test to confirm that there was no statistically significant difference in MBIE scores between non-strike tutorial groups and tutorial groups who striked (p= 0.8088). This is perhaps because students affected by strikes often had very candid conversations with their tutors about why strikes were happening and why tutors may have felt the need to participate in picket lines. Exclusion may have been felt because tutorial sessions were not held and tutors could be seen as placing their own needs over the needs of their students. Yet, inclusion may have been felt because of tutors’ communication of their, often, very personal reasons for striking.

I collected student demographics only during Survey #1 to minimise the amount of time participants had to spend filling out the surveys. For this reason, if a student did not fill out Survey #1, but did fill out Survey #2, their demographic data were missing. I personally reached out to each of these students to ask them to fill out a separate demographic survey. In total, only 10 students never filled in demographic information
and were disregarded from analysis. This left 156 unique student participants in total out of the 249 students who originally signed up for the study. 148 unique student responses were collected for Survey #1, while 105 unique student responses were submitted for Survey #2. Respondents were given the option to opt out of certain questions, for example with the question asking about their sexual orientation.

Demographics for Survey #1 respondents and Survey #2 respondents were very similar, despite the drop in the number of Survey #2 respondents. With that, the mean age for Survey #1 was 18.72 and 18.81 for Survey #2. This makes sense as the majority of enrolled students in the two courses from School A and School B were first year students. This translated to students’ self-reported school years too where 80.41% of Survey #1 respondents were first year students, and 81.90% of Survey #2 were first year students. Both surveys also had a similar gender breakdown. For Survey #1, 77.03% of respondents were female, 20.27% male, and 2.70% were non-binary or transgender. For Survey #2, 80.95% of respondents were female, 15.24% were male, and 3.81% were non-binary or transgender. These figures are perhaps influenced by School B’s gender breakdown with 67.7% female entrants in 2021/2022. Gender reports for School A or per programme were unavailable. Just about 75% of respondents were heterosexual for both surveys. 20% were non-heterosexual in Survey #2 and 21% were non-heterosexual in Survey #1. Generally, about half of the student respondents reported being involved in at least one university activity for both surveys. This is important because if one survey group showed higher levels of involvement, that may have been associated with higher levels of inclusion as they may know more students personally in their courses.

Both surveys also had similar proportions of students with disabilities and BAME students. Survey #1 had 7.43% of students reporting a disability, while Survey #2 had 11.43% of students reporting a disability. This is lower than the university’s report of 17.10% of all undergraduates having a disability. Based on these figures, 92.57% and 88.57% of students in Survey #1 and Survey #2, respectively, had no disability reported. For Survey #1, 62.16% of students were white, 31.76% were Asian, 3.38% were Middle Eastern, 2.03% were Black, and 0.68% were Latinx. For Survey #2, a similar breakdown was reported whereby 59.05% of students were white, 34.29% were Asian, 2.86% were Middle Eastern, 2.86% were Black, and 0.95% were Latinx. This is aligned with the percentage of white incoming undergraduates in the 2021-2022 academic year which was about 64%. The sample does over index on Asian students though as 14.16% of incoming undergraduates in 2021-2022 identified as Asian. This may simply be an over index as enrolments of Asian students may be higher in the two schools chosen for this study, than in the overall university population. Unfortunately, this information was unavailable. It is also important to note that of the students who self-identified as white, 58.06% were British across the whole participant sample. The next largest ethnicity was Chinese with 71.15% of Asian students self-identifying as Chinese.
The ordinal nature of my data created implications for analysis. The main implication was that differences between scores were difficult to comment on as, for instance, the distance between *completely agree* to *moderately disagree* cannot be assumed to be equivalent to the difference between *moderately agree* to *slightly agree*. Instead, I discuss the scores as definitively higher or lower. Considering the ordinal nature of the data, the first step in conducting statistical analysis included utilising parametric techniques to check for any correlation between overall MBIE scores and students’ reported demographics. Depending on the demographic group being analysed, this was either done using a Welch two-sample T-test or a two-way ANOVA test. The null hypothesis in this case meant there was no relationship between being in a certain demographic group and having a higher MBIE composite score. A t-test was used to specifically understand if the null hypothesis could be rejected for disability, university involvement, sexuality, and race and ethnicity when combined into white and BAME demographic groups. To get a composite MBIE score for each student, the scores from each of the 11 MBIE survey items were added up. Given *Item #6- Last to Know* in the MBIE survey was a negatively worded statement, the inverse of the scores were coded for prior to the addition into a composite score. If a student put 1 for *Item #6- Last to Know*, it was coded as a 6. If a student put 6 for *Item #6- Last to Know*, then it was coded as a 1. Lastly, for the composite score changes, I used a Chi-Squared test to understand if certain groups experienced higher instances of increasing, or decreasing, scores than would be expected.

In this analysis, I also checked the statistical significance of variance by demographic groups in survey item scores. This required either a Mann-Whitney Wilcoxon analysis or Kruskal-Wallis Chi-Squared test. For the same reasons as discussed above, these two tests are appropriate as the ordinal data requires non-parametric statistical techniques. For disability, university involvement, and race and ethnicity (white and BAME), I used a Mann-Whitney Wilcoxon analysis. For school year, gender, and the other two race and ethnicity groupings, I used the Kruskal-Wallis Chi-Squared test given there were more than two categories for these demographic groups. The last analysis I discuss in this section involved investigating how different demographic groups scored on average for certain items that changed significantly from Survey #1 to Survey #2. To conduct this analysis, I used additional ANOVA testing to compute an analysis of variance table with a linear model input. I did this to understand the influence of the most salient demographics on the scores for certain survey items, while adjusting and controlling for each variable.

### 6.2.1: Aggregated Scale Analysis
As seen in Table 6.3, the null hypothesis could not be rejected at $p = 0.05$ for any of these demographic groups in either survey. In this way, there is weak evidence of a relationship between being in a certain demographic group and having a higher MBIE composite score (perception of inclusion) for either of the two surveys. A t-test was used for the majority of this analysis, while a two-way ANOVA was used for statistical analysis of students’ years in school, gender, as well as race and ethnicity too. In the case of race and ethnicity, two ANOVA analyses were run in addition to the aforementioned white and BAME t-test. Race and ethnicity were also split into white, Asian, and other (Latinx, Black, Middle Eastern), as well as white, Asian, mixed, and other (Latinx, Black, and Middle Eastern). The reason three separate analyses were run for race and ethnicity is because there are various ways to group students by their reported race and ethnicities. For one, Latinx, Black, and Middle Eastern students were too few in number thus needed to be grouped together for statistical power. At the same time, I wanted to check if students who self-reported as being mixed had a unique intersectional experience in comparison to if I simply grouped them with their BAME demographic groups. Being mixed myself, I know that sometimes being mixed is salient in certain contexts, whereas other times, my whiteness or my Latinx identity may feel more relevant. On the contrary, statistical analysis showed that none of these racial and ethnic demographic groupings had a statistically significant relationship with higher composite MBIE scores. The same was true for statistical analysis of gender and students’ years in school.
<table>
<thead>
<tr>
<th>Demographic Groups</th>
<th>Survey #1</th>
<th>Survey #2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>n</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>46.27</td>
<td>114</td>
</tr>
<tr>
<td>Male</td>
<td>47.03</td>
<td>30</td>
</tr>
<tr>
<td>Non-binary/Transgender</td>
<td>47.50</td>
<td>4</td>
</tr>
<tr>
<td>Sexuality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-heterosexual</td>
<td>43.58</td>
<td>31</td>
</tr>
<tr>
<td>Heterosexual</td>
<td>46.99</td>
<td>111</td>
</tr>
<tr>
<td>Disability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Students with disabilities</td>
<td>50.00</td>
<td>11</td>
</tr>
<tr>
<td>Students with no disabilities</td>
<td>46.17</td>
<td>137</td>
</tr>
<tr>
<td>Year in School</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First year</td>
<td>47.09</td>
<td>119</td>
</tr>
<tr>
<td>Second year</td>
<td>42.15</td>
<td>20</td>
</tr>
<tr>
<td>Third year</td>
<td>47.43</td>
<td>7</td>
</tr>
<tr>
<td>Fourth year</td>
<td>48.50</td>
<td>2</td>
</tr>
<tr>
<td>School</td>
<td></td>
<td></td>
</tr>
<tr>
<td>School A</td>
<td>47.06</td>
<td>48</td>
</tr>
<tr>
<td>School B</td>
<td>46.16</td>
<td>100</td>
</tr>
<tr>
<td>University Involvement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Involved in University Activities</td>
<td>46.24</td>
<td>79</td>
</tr>
<tr>
<td>Not Involved in University Activities</td>
<td>46.71</td>
<td>69</td>
</tr>
<tr>
<td>Race &amp; Ethnicity (Binary)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>45.97</td>
<td>92</td>
</tr>
<tr>
<td>BAME</td>
<td>47.27</td>
<td>56</td>
</tr>
<tr>
<td>Race &amp; Ethnicity (not including Mixed)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>45.97</td>
<td>92</td>
</tr>
<tr>
<td>Asian</td>
<td>47.02</td>
<td>47</td>
</tr>
<tr>
<td>Other</td>
<td>48.56</td>
<td>9</td>
</tr>
<tr>
<td>Race &amp; Ethnicity (including Mixed)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>45.97</td>
<td>92</td>
</tr>
<tr>
<td>Asian</td>
<td>47.59</td>
<td>42</td>
</tr>
<tr>
<td>Mixed</td>
<td>48.57</td>
<td>7</td>
</tr>
<tr>
<td>Other</td>
<td>44.14</td>
<td>7</td>
</tr>
</tbody>
</table>
A key finding included when not considering demographics, the overall cumulative MBIE score had a significantly higher average of 46.30 at Survey #2 than compared to 43.66 at Survey #1 (p=0.0143). Table 6.3 additionally shows how the average cumulative MBIE score was higher in Survey #2 for all demographic groups except third year students and students identifying as racially mixed. Even so, average perceived inclusion overall was higher in Survey #2, with 66 being the maximum MBIE cumulative score. At the same time, it is important to consider how the composite MBIE scores differed for people who took both surveys (n= 95). Of those, 55.79% of students had higher composite MBIE scores in Survey #2, while 36.84% and 7.37% reported lower scores or no difference, respectively. Utilising a Chi-squared test, the only demographic that had a statistically significant difference of observed frequencies than what would be expected was sexuality. At Survey #2, 10.00% of non-heterosexual students had lower composite MBIE scores than in Survey #1, while 43.66% of scores for heterosexual students were lower for the same time period (p= 0.0146). 89.47% of non-heterosexual students had higher composite MBIE scores at Survey #2 and 49.30% of heterosexual students experienced the same. This is reflected in Table 6.3, where there is weak evidence (p= 0.08) for a statistically significant difference between heterosexual and non-heterosexual students’ cumulative scores on Survey #1.

As overall cumulative MBIE scores were statistically higher at Survey #2 in comparison to Survey #1, it is important to understand which survey items saw the largest increases in average scores and which changed less drastically. The average response scores for 10 of the 11 survey items increased from Survey #1 to Survey #2. The only survey item that was lower in average score in Survey #2 was Item #6- Last to Know, as detailed in Figure 6.4. Students scored on average 4.90 in Survey #1 for this item, but was lower at a 4.82 average in Survey #2. The context of the strikes are likely to blame for this, as they resulted in quickly changing timetables and cancellations largely out of direct control of the tutor. The two survey items that witnessed the largest increase in average scores between Survey #1 and Survey #2 were Item #4- Invited to Informal Activities and Item #3- Informed About Informal Activities, which may be seen in Figure 6.5 and Figure 6.6. Item #4- Invited to Informal Activities had an average score of 2.76 on Survey #1, but was higher at 3.42 in Survey #2. Similarly, students averaged 3.35 on Item #3- Informed About Informal Activities on Survey #1 and the average was higher at 3.77 by Survey #2. This likely is a result of students simply getting to know each other in the classroom more over time. By Survey #2, students would have been working together in small groups for several weeks, as well as exchanging phone numbers and personal details. It is unsurprising that these two survey items saw higher scores on Survey #2 than on Survey #1 given their focus on peer interactions. A full overview of student responses to each survey item per survey may be seen in Appendix C.5.
Figure 6.4: Likert Scale Responses per Survey for Survey Item #6- Last to Know

Figure 6.5: Likert Scale Responses per Survey for Survey Item #4- Invited to Informal Activities

Figure 6.6: Likert Scale Responses per Survey for Survey Item #3- Informed About Informal Activities
6.2.2: Demographic Patterns of Survey #1
<table>
<thead>
<tr>
<th>MBIE Survey Item</th>
<th>Gender</th>
<th>Sexuality</th>
<th>Disability</th>
<th>Year</th>
<th>School</th>
<th>University Involvement</th>
<th>Race &amp; Ethnicity: White</th>
<th>Race &amp; Ethnicity: White, Asian, Other</th>
<th>Race &amp; Ethnicity: White, Asian, Mixed, Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. My peers in my tutorial group openly share information with me.</td>
<td>0.9807</td>
<td>0.7029</td>
<td>0.5289</td>
<td>0.2974</td>
<td>0.6207</td>
<td>0.0451*</td>
<td>0.0053*</td>
<td>0.0202*</td>
<td>0.0090*</td>
</tr>
<tr>
<td>2. I am typically involved and invited to actively participate in activities in my peers.</td>
<td>0.7758</td>
<td>0.8721</td>
<td>0.2204</td>
<td>0.3298</td>
<td>0.2232</td>
<td>0.4512</td>
<td>0.2912</td>
<td>0.5560</td>
<td>0.7696</td>
</tr>
<tr>
<td>3. I am always informed about informal social activities and events with my peers in my tutorial group.</td>
<td>0.4161</td>
<td>0.9980</td>
<td>0.2380</td>
<td>0.0428*</td>
<td>0.9105</td>
<td>0.1551</td>
<td>0.1841</td>
<td>0.1252</td>
<td>0.3631</td>
</tr>
<tr>
<td>4. I am often invited to join my peers from my tutorial group when they go out for a meal, drinks, or a study session.</td>
<td>0.5949</td>
<td>0.4605</td>
<td>0.2809</td>
<td>0.4397</td>
<td>0.0789</td>
<td>0.3466</td>
<td>0.1350</td>
<td>0.2920</td>
<td>0.3539</td>
</tr>
<tr>
<td>5. I am often invited to contribute my opinion in tutorials by my tutor.</td>
<td>0.7873</td>
<td>0.2297</td>
<td>0.8493</td>
<td>0.6360</td>
<td>0.7461</td>
<td>0.0307*</td>
<td>0.1067</td>
<td>0.0153*</td>
<td>0.102</td>
</tr>
<tr>
<td>6. I am usually among the last to know about important changes in the tutorial.</td>
<td>0.8341</td>
<td>0.8957</td>
<td>0.9876</td>
<td>0.5105</td>
<td>0.1530</td>
<td>0.0840</td>
<td>0.4489</td>
<td>0.0412*</td>
<td>0.0360*</td>
</tr>
<tr>
<td>7. My tutor often asks for my opinion before making important decisions.</td>
<td>0.8478</td>
<td>0.0326*</td>
<td>0.2157</td>
<td>0.0415*</td>
<td>0.1657</td>
<td>0.5438</td>
<td>0.0135*</td>
<td>0.0469*</td>
<td>0.0040*</td>
</tr>
<tr>
<td>8. I frequently receive communication from my tutor.</td>
<td>0.7260</td>
<td>0.1194</td>
<td>0.1227</td>
<td>0.2142</td>
<td>0.3268</td>
<td>0.7264</td>
<td>0.8945</td>
<td>0.1028</td>
<td>0.5956</td>
</tr>
<tr>
<td>9. I am invited to actively participate in communication with my tutor.</td>
<td>0.5127</td>
<td>0.0661</td>
<td>0.4884</td>
<td>0.6535</td>
<td>0.7665</td>
<td>0.0511</td>
<td>0.5785</td>
<td>0.0780</td>
<td>0.3757</td>
</tr>
<tr>
<td>10. I am able to influence decisions that affect my tutorial.</td>
<td>0.1819</td>
<td>0.0930</td>
<td>0.2523</td>
<td>0.0136*</td>
<td>0.6683</td>
<td>0.2698</td>
<td>0.7667</td>
<td>0.9327</td>
<td>0.7194</td>
</tr>
<tr>
<td>11. When there is a meeting with the course organisers outside of my tutorial, I am invited to participate by my tutor.</td>
<td>0.7380</td>
<td>0.1005</td>
<td>0.2296</td>
<td>0.8186</td>
<td>0.3452</td>
<td>0.1992</td>
<td>0.0885</td>
<td>0.0874</td>
<td>0.0833</td>
</tr>
</tbody>
</table>

* = significant at p-value < 0.05
Even though there were no significant differences in composite MBIE scores between different demographic groups, there were significant differences in individual survey item scores for certain demographic groups in Survey #1. These differences are detailed in Table 6.7. Six survey items in Survey #1 had variance between demographic groups that was statistically significant, including for racial and ethnic groups, year, involvement in university activities, and sexuality. No variance in the item responses was statistically significant in analysis done by gender, disability, or school (School A or School B). To check whether any of the statistically significant demographic groups remained significant when controlling for each other, I ran a two-way ANOVA test. Interestingly, sexuality was the only demographic group that remained significant (p=0.0472) when controlled for the other variables, which was likely driven by students’ responses to Item #7- Opinion Before Decisions. Students who identified as non-heterosexual scored an average of 3.65 on Item #7- Opinion Before Decisions which is lower than heterosexual students’ average score of 4.26 (p=0.0326). This is in opposition to tutors reporting an average score of 5.14 for the tutor equivalent survey item showing that tutors overall did feel they asked for student opinions ahead of making important decisions.

Variance in scores by university involvement also proved significant. Students who were involved in university activities scored 4.27 on average, while students not involved scored 4.59 on average (p= 0.0451) for Item #1- Share Information. Involved students also scored 5.39 on average as students not involved scored an average of 5.06 for Item #5- Contribute Opinion. Therefore, students involved in university activities perhaps feel a slightly closer relationship with their tutor than students not involved. Conversely though, those same students may feel a little more disconnected from their tutorial peers due to their likely more solidified peer relationships outside of their tutorial groups through their university involvement. This could be further explained by the 3.67/5.00 average score of trust in tutorials by students involved in university activities, in comparison to the 4.00/5.00 students not involved felt in terms of tutorial trust (p=0.0163). Van Gijn-Grosvenor & Huisman (2020) similarly found that students reported involvement with university clubs and societies heightened students’ perceptions of belonging at university. This further indicates that there is a chance that students may feel included in one context of university, but that may make them feel less included in another.

The academic year students were in was also significant for three questions on Survey #1. This included Item #10- Influence Decisions (p=0.0136), Item #3- Informed About Informal Activities (p=0.0428), and Item #7- Opinion Before Decisions (p=0.0415). For Item #10- Informed About Informal Activities and Item #7- Opinion Before Decisions, students in Year 2 had the lowest average scores with 3.35 and 3.3 on average, respectively. First years, who made up the majority of the tutorial groups, scored on
average 4.16 and 4.23 for these two survey items, respectively. Both of these survey items had to do with tutor communication showing that tutors may have instead been focusing and tailoring their communications to the Year 1 students. Yet, Year 3 and Year 4 students actually scored higher on average for both of these statements than first year students. There is a chance though that these students were exchange students or simply taking the introductory course as a filler in their final years. Consequently, they may have had less expectations for tutor communication than the Year 1 or Year 2 students. Year 3 students scored the lowest with 2.29 on average (p=0.0428) for Item #3- Informed About Informal Activities. There was only a 0.01 difference between the average score for Year 1 and Year 4 students, who each scored on average 3.51 and 3.50. Year 4 students may have felt more secure in not being invited to tutorial group social activities given their distance from Year 1, whereas Year 3 students may have felt a little more inclined to socialise with the first years. Even though Year 3 students felt a lack of communication about informal social activities, they actually had the highest average score for perceptions of community in tutorial groups. Year 3 students maintained an average community score of 4.43/5.00, whereas first year students averaged 3.82/5.00 (p=0.0294). This shows that despite the feelings that they were not being communicated with from other peers, Year 3 students still felt a relatively high sense of community in their tutorial groups. For first years, this hints at social anxiety in their tutorial groups whereby they felt peers did not communicate with them about social activities and they lacked as high of a sense of trust as Year 3 students.

While variance by a student’s year was significant for three MBIE survey items, variance by race and ethnicity was significant for four of the 11 MBIE survey items. Given race and ethnicity proved to be significantly related to variance for the most MBIE survey items, Figures 6.8 through Figure 6.11 detail each racial group’s scores for the four survey items. The four MBIE survey items that were significant by race and ethnicity were Item #1- Share Information, Item #5- Contribute Opinion, Item #6- Last to Know, and Item #7- Opinion Before Decisions. Only Item #1- Share Information has to do with students’ peers, while Items #5-#7 have to do with tutor communication. Related to this, it is worth noting that analysis by grouping students into white, Asian, and a catch-all “other” category proved most significant. Only two survey items were significant when analysis was done grouping students into white and BAME groups. Adding a mixed category into the analysis revealed just three significant MBIE survey items.

As detailed in Figure 6.8 and Figure 6.11, my analysis found white students were less likely than Asian students to agree with Item #1- Share Information and Item #7- Opinion Before Decisions. For these two survey items, white students averaged 4.22 (p=0.0202) and 3.95 (p=0.0469), respectively. Asian students reported average scores of 4.74 and 4.44 for these two survey items. According to the literature, BAME students
can often struggle with social integration, reporting difficulties in conversing with course peers who do not identify as a minority and experiencing the white-centric classroom discourse as exclusionary (Arday et al., 2022; Kauser et al., 2021). The findings detailed in Figure 6.8 and Figure 6.11 contradict the literature in this case. For Item #7- Opinion Before Decisions, white students seem to expect more invitations from the tutor ahead of important decisions than Asian students. This could be connected to the lower average perceptions of community by white students in comparison to Asian and students of other racial and ethnic backgrounds. The average perception of community score for white students was 3.61, while it was 3.96 for Asian students and 4.33 for all other students on a scale of 1-5 (p= 0.0100). As in Figure 6.9, Asian students scored the lowest on Item #5- Contribute Opinion with 4.83 (p=0.0153), showing that perhaps there is some expectation that tutors will directly ask students to contribute to the tutorial discussion. This is generally not the case though because tutors will not pick on students to contribute, but rather open up the discussion floor for anyone willing to jump in. Interestingly, for Item #6- Last to Know, displayed in Figure 6.10, white students’ scores were on average 2.01, while Asian students scored an average of 2.41 and BAME and non-Asian students averaged 1.44. This more generally shows that tutors do communicate important changes to students, but that additional communication may be needed for Asian students to feel more adequately informed.

Figure 6.8: Likert Scale Responses per Racial Group for Survey #1 Item #1- Share Information

![Likert Scale Responses per Racial Group for Survey #1 Item #1- Share Information](image)
Figure 6.9: Likert Scale Responses per Racial Group for Survey #1 Item #5 - Contribute Opinion

Figure 6.10: Likert Scale Responses per Racial Group for Survey #1 Item #6 - Last to Know

Figure 6.11: Likert Scale Responses per Racial Group for Survey #1 Item #7 - Opinion Before Decisions
Overall, Item #7- Opinion Before Decisions proved to be most significant across demographic groups when it comes to variance, followed by Item #1- Share Information. Item #7- Opinion Before Decisions was significant in five different demographic analyses, while Item #1- Share Information was significant in four analyses. The difference between them though is that Item #7- Opinion Before Decisions addressed tutor communication, while Item #1- Share Information addressed peer communication. Item #7- Opinion Before Decisions also explores students’ perception of how much tutors are communicating ahead of decision making. This prompts the question of what students’ expectations of tutors’ responsibilities are, and what cadence tutors are communicating with students about decisions that are beyond their paygrade. Item #1- Share Information was more difficult for tutors to address, and likely depended more directly on the overall tutorial dynamics. In fact, in Survey #1, of the six survey items flagged in analysis, just two had to do with peer interactions compared to four that had to do with tutors’ actions. This shows significant room for tutors to improve communication to foster perceptions of inclusion for students regardless of identity, especially given tutors felt their communication ahead of important decisions was high. In turn, this could help feed back into boosting peer-to-peer communication and feelings of inclusion.

6.2.3: Demographic Patterns of Survey #2
### Table 6.12: Significance of Demographics per Survey #2 MBIE Item

<table>
<thead>
<tr>
<th>MBIE Survey Item</th>
<th>Gender</th>
<th>Sexuality</th>
<th>Disability</th>
<th>Year</th>
<th>School</th>
<th>University Involvement</th>
<th>Race &amp; Ethnicity: White, BAME</th>
<th>Race &amp; Ethnicity: White, Asian, Other</th>
<th>Race &amp; Ethnicity: White, Asian, Mixed, Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. My peers in my tutorial group openly share information with me.</td>
<td>0.8425</td>
<td>0.9203</td>
<td>0.9698</td>
<td>0.1043</td>
<td>0.5134</td>
<td>0.1121</td>
<td>0.0193*</td>
<td>0.0324*</td>
<td>0.0617</td>
</tr>
<tr>
<td>2. I am typically involved and invited to actively participate in activities in my tutorial group by my peers.</td>
<td>0.7620</td>
<td>0.2292</td>
<td>0.8216</td>
<td>0.2775</td>
<td>0.2015</td>
<td>0.4084</td>
<td>0.6664</td>
<td>0.7843</td>
<td>0.8658</td>
</tr>
<tr>
<td>3. I am always informed about informal social activities and events with my peers in my tutorial group.</td>
<td>0.1011</td>
<td>0.7013</td>
<td>0.8505</td>
<td>0.0546</td>
<td>0.5630</td>
<td>0.4578</td>
<td>0.4603</td>
<td>0.4688</td>
<td>0.7241</td>
</tr>
<tr>
<td>4. I am often invited to join my peers from my tutorial group when they go out for a meal, drinks, or a study session.</td>
<td>0.5641</td>
<td>0.4924</td>
<td>0.7501</td>
<td>0.0797</td>
<td>0.7213</td>
<td>0.2688</td>
<td>0.7229</td>
<td>0.6544</td>
<td>0.6367</td>
</tr>
<tr>
<td>5. I am often invited to contribute my opinion in tutorials by my tutor.</td>
<td>0.1068</td>
<td>0.1422</td>
<td>0.2896</td>
<td>0.2991</td>
<td>0.6117</td>
<td>0.8835</td>
<td>0.7024</td>
<td>0.0458*</td>
<td>0.2363</td>
</tr>
<tr>
<td>6. I am usually among the last to know about important changes in the tutorial.</td>
<td>0.7657</td>
<td>0.0227*</td>
<td>0.0470*</td>
<td>0.0645</td>
<td>0.4398</td>
<td>0.1398</td>
<td>0.0145*</td>
<td>0.0198*</td>
<td>0.0063*</td>
</tr>
<tr>
<td>7. My tutor often asks for my opinion before making important decisions.</td>
<td>0.8077</td>
<td>0.1718</td>
<td>0.2281</td>
<td>0.0319*</td>
<td>0.0288*</td>
<td>0.0747</td>
<td>0.0531</td>
<td>0.1367</td>
<td>0.0816</td>
</tr>
<tr>
<td>8. I frequently receive communication from my tutor.</td>
<td>0.2681</td>
<td>0.9710</td>
<td>0.1208</td>
<td>0.2181</td>
<td>0.1327</td>
<td>0.9226</td>
<td>0.8905</td>
<td>0.5288</td>
<td>0.3732</td>
</tr>
<tr>
<td>9. I am invited to actively participate in communication with my tutor.</td>
<td>0.6647</td>
<td>0.4998</td>
<td>0.1035</td>
<td>0.0950</td>
<td>0.3213</td>
<td>0.9711</td>
<td>0.1732</td>
<td>0.0784</td>
<td>0.3353</td>
</tr>
<tr>
<td>10. I am able to influence decisions that affect my tutorial.</td>
<td>0.3825</td>
<td>0.2400</td>
<td>0.5293</td>
<td>0.0680</td>
<td>0.2795</td>
<td>0.8783</td>
<td>0.1620</td>
<td>0.3741</td>
<td>0.4131</td>
</tr>
<tr>
<td>11. When there is a meeting with the course organisers outside of my tutorial, I am invited to participate by my tutor.</td>
<td>0.6018</td>
<td>0.2077</td>
<td>0.1237</td>
<td>0.2815</td>
<td>0.8890</td>
<td>0.0417*</td>
<td>0.0216*</td>
<td>0.0313*</td>
<td>0.0712</td>
</tr>
</tbody>
</table>

* = significant at p-value < 0.05
The overall number of survey items triggered as significant in terms of demographic groups was lower in Survey #2 than in Survey #2, as per Table 6.12. Five survey items’ variance were statistically significant by demographics compared to six survey items in Survey #1. This can be considered a small sign that inclusion for students was stronger at Survey #2 than earlier in the term at Survey #1. At the same time, variance in item-by-item scores for more demographic groups proved significant. With that, gender was the only demographic analysis that had no significance in terms of individual MBIE survey item’s scoring variance. Gender was also insignificant for Survey #1. Just as with Survey #1, I also checked whether any of the statistically significant demographic groups remained significant when controlled for each other. In running a two-way ANOVA test, no demographic groups remained significant when variables were controlled.

Analysis by disability status, sexuality, school year, school, and university involvement was significant for one MBIE survey item each. Once again, students who were involved with at least one university activity felt more excluded from meetings with course organisers than students who were not involved with university activities. Those who were involved averaged 3.20 for Item #11- Outside Meeting Invites, compared to the average of 3.83 for students who were not involved in university activities (p=0.0417). This is similar to Survey #1 where involved students felt lower levels of peer information sharing and invitations from tutors to share their opinions. Yet, it is important to keep in mind that students involved in university activities placed less importance on feeling community in their tutorial groups compared to those not involved in university activities (p=0.0218). For Item #6- Last to Know, students without disabilities and heterosexual students both felt less ‘in the know’ respectively (p=0.0470 and p=0.0227). Students with disabilities scored an average 5.25, while students without disabilities averaged 4.80. Just the same, heterosexual students averaged 4.73, while non-heterosexual students averaged 5.38 for the same survey item. A lack of tutor communication was also felt once again by students in their second year where students felt less welcome to give their opinion ahead of important decisions (p= 0.0319). Students in Year 2 averaged 3.36 for Item #7- Opinion Before Decisions, while students in Year 1 averaged 4.58, students in Year 3 averaged 4.33, and fourth year students averaged 5.00. The same sentiment was felt by students in School B rather than students in School A. This is perhaps linked to the frequency of strikes in School B. As discussed, it is rare for tutors to consult the opinion of their students on whether they should strike or not. More frequently, tutors would approach strike discussions with students already having made the decision to strike or not to strike. This may be linked to why students in School B averaged 4.15 compared to School A’s average of 4.76 for Item #7- Opinion Before Decisions.
Similar to Survey #1, analysis by race and ethnicity for Survey #2 questionnaire items was most significant when done for the three groups: white, Asian, and an all-encompassing category for everyone else. The variance for four survey items was significant when analysis was done in this way, whereas only three survey items were significant when analysis was done by grouping students into white and BAME. Just one survey item was significant when a mixed category was added to the analysis. The variance in scores for these four survey items may be seen in Figure 6.13 through Figure 6.16. Notably, for Item #1- Share Information and Item #11- Outside Meeting Invites, white students had the lowest average scores of 4.44 and 3.21 each (p=0.0324 and p=0.0216). For Item #1- Share Information, significantly more students who were BAME and non-Asian reported completely agreeing and moderately agreeing, while white students reported more slightly agreeing that their tutorial group openly shares information with them. For Item #11- Outside Meeting Invites, it is also notable that students who did not identify as white nor Asian scored significantly lower too with an average score of 3.29, compared to white students’ average score of 3.21 and Asian students’ 4.09 (p=0.0216). However, Asian students scored the lowest on average with 5.23 and 4.31 on Item #5- Contribute Opinion (p=0.0458) and on Item #6- Last to Know (p=0.0198). Both of these patterns were seen in Survey #1 too, and notably for Item #6- Last to Know, Asian students’ scores worsened by -0.28 over time. This again shows how tutors can perhaps change the way they communicate with Asian students to ensure all students regardless of identity feel invited to contribute and like they are not being left out of communications. White students, on the other hand, displayed a desire for more invitations to engage with course organisers through their tutors and for more communication with peers.

Figure 6.13: Likert Scale Responses per Racial Group for Survey #2 Item #1- Share Information

![Graph showing Likert Scale responses per racial group for Survey #2 Item #1- Share Information.](image-url)
Figure 6.14: Likert Scale Responses per Racial Group for Survey #2 Item #5- Contribute Opinion

Figure 6.15: Likert Scale Responses per Racial Group for Survey #2 Item #6- Last to Know

Figure 6.16: Likert Scale Responses per Racial Group for Survey #2 Item #11- Outside Meeting Invites
Comparing Survey #1 and Survey #2 hints at a few key differences between student sentiments of inclusion at each timepoint. In Survey #1, variances in Item #1-Share Information, Item #3- Informed About Informal Activities, Item #5- Contribute Opinion, Item #6- Last to Know, Item #7- Opinion Before Decisions, and Item #10-Influence Decisions were significant based on student demographics. Survey #2 had just five survey items with significant differences. Survey #2 did not have Item #3-Informed About Informal Activities and Item #10-Influence Decisions, but did see the addition of Item #11- Outside Meeting Invites. The loss of Item #3- Informed About Informal Activities as significant shows that maybe Year 3 students were able to overcome the feelings of being disconnected socially from their peers. This is keeping in mind that their peers were namely first year students. Similarly, Year 2 students who in Survey #1 felt less able to influence decisions may have built a rapport with their tutors over time. Thus, in Survey #2, they felt they could more easily influence decisions than at the beginning of the term where they entered at a disadvantage as a minority in terms of school year. It remains concerning though that Year 2 students had the lowest perception of community in their tutorials at 3.64/5.00 compared to 4.12/5.00 for first year students (p=0.0207).

Moreover, the addition of Item #11- Outside Meeting Invites actually makes sense and is to be expected to an extent. Meetings with the course organiser are less likely to occur during the start of term. In contrast, closer to the end of term, course organisers may schedule meetings to act as resources for final essays or presentations. Also, due to the UCU strikes that happened towards the end of the term, it is likely that course organisers were more involved than in the average term. As a result, it is perhaps the case that there were more opportunities to meet with course organisers towards the end of the term when Survey #2 was open, but some students (namely white students) felt less knowledgeable of when those meetings were. To back this up, tutors rated their communication with students about course organiser meetings as a 2.00 on average. This shows that perhaps tutors are aware of the disconnect between students and course organisers.

Interestingly, of the five survey items which had significant variance, four had to do with the tutor themselves, of which three explicitly referred to the tutor. Just one survey item with variance significantly related to demographics focused on peers. This goes to show that by and large, tutors have the opportunity to alter their behaviours to more effectively foster inclusion for all students. For instance, the survey item which had the most significant variance by demographic groups was Item #6- Last to Know. Students who identified as heterosexual, as not having a disability, and as Asian, all scored the lowest on average for this survey item. This statement does not explicitly call out tutors, but it does imply a lack of or inconsistent tutor communication.
6.3: Discussion

This study provides insight into whether certain demographic groups feel excluded in their tutorial sessions during a time when diversity and inclusion are at the forefront of higher education discourse. At a high-level and taking into account each group’s cumulative Mor Barak Inclusion-Exclusion scores, no group had significantly lower composite MBIE scores on either of the two surveys. As a result, this study lacks evidence of an overt inequality regime in place within the tutorials studied when analysed in terms of sexuality, disability, gender, race and ethnicity, school year, university involvement, or school. This includes when students are analysed through the more intersectional lens of analysis distinguishing mixed students. As someone who identifies as mixed, I do tend to believe the mixed identity feels relevant to my experiences of inclusion/exclusion in the classroom context. On the contrary, the findings show that this was not the case. In fact, the results point to the tutorials being, on average, overall slightly to moderately inclusive by end of term. While a direct link to the effects of the inclusion training cannot be drawn, there is some evidence to believe that the inclusion trainings can continue to be refined even more.

A main finding of this work also includes evidence that overall MBIE scores were significantly higher in Survey #2 with an average of 46.30 in comparison to 43.66 in Survey #1 (p=0.0143). Accordingly, student perceptions of inclusion were higher in Survey #2 even as tutor inclusion stayed about the same in both surveys (decreasing only -0.02). Given there was no control group, it cannot be definitively concluded that the trainings were a reason for the higher perceptions of inclusion in Survey #2 than in Survey #1. Even so, this is the first study, to my knowledge, to conduct a demographic analysis of a wide-scale inclusion survey for undergraduate students in more than one school focusing on the classroom setting. While this finding is worthwhile both empirically and methodologically, it would be interesting in the future to expand the study’s scope to understand if the students feel different levels of inclusion in each of their tutorial sessions or courses they are enrolled in. It would also be beneficial to conduct the training with a control group to create a more direct link between the inclusion trainings administered and perceptions of inclusion.

While it is difficult to situate the overall change in MBIE scores and item-by-item changes in the literature, given this is the first study to adapt the MBIE Scale to the classroom, it is still important to discuss both sets of findings. One key finding was that 55.79% of the cumulative MBIE scores for students who took both surveys increased. An interesting caveat to this was that heterosexual students reported a decrease in overall perceptions of inclusion 43.66% of the time, compared to non-heterosexual students reporting a decrease only 10% of the time. This contrasts the literature, which expects more perceptions of exclusion for non-heterosexual students more than for
heterosexual students (Ferfolja et al., 2020). The reasons for this finding are unclear, but may be due to tutors’ priming of trying to include those from minority backgrounds. When the surveys was analysed item-by-item, it was also notable that Item #3- Informed About Informal Activities and Item #4- Invited to Informal Activities both had the largest increases in average scores. This is aligned with the findings of Rienties et al. (2013) who found that most students will experience growth in their friendships within the classroom overtime.

Analysing the inclusion survey datasets item-by-item do show that some demographic groups feel lower levels of inclusion in certain aspects of their tutorial experience. This contribution highlights that certain demographic groups report that lower levels of inclusion are felt not necessarily because of peer interactions, but due to interactions with the tutors. This is true in both Survey #1 and Survey #2. Overall, the survey items that had significant variance by demographic groups included Item #1- Share Information, Item #3- Informed About Informal Activities, Item #5- Contribute Opinion, Item #6- Last to Know, Item #7- Opinion Before Decisions, Item #10- Influence Decisions, and Item #11- Outside Meeting Invites. Of these survey items, two have to do with peer interactions and five have to do with students’ tutors. More specifically, all five of the tutor survey items have to do with communication from the tutor to the students, but not about the course content. Rather, these survey items involved invitations for student opinions, decision making, and full course staff communication.

This finding of the influence of tutors on perceptions of inclusion/exclusion is aligned with and contributes to existing literature. Finkelstein et al. (2021) found in a scoping review of inclusive teaching studies that there is more to inclusive teaching than collaboration and communication with students. Rather, inclusive teaching can also be framed in terms of organisational practices whereby teaching staff help students establish routines and facilitate transitions. Collaboration must also include helping students work with teaching teams. Both of these aspects of inclusive teaching are incredibly relevant to this study and reflected in the MBIE survey items with significant variance. Most notably, it can be inferred that perhaps students needed more communication and deliberation with students on the strikes that occurred. Furthermore, perhaps tutors may have put in extra effort to make the course organisers more visible given tutorials were being cancelled. Tutors, to an extent, may already be aware of this given their low average responses to questions regarding communication to students about course organisers. In this way, certain demographic groups may have felt more included in terms of decision making, invitations for their opinions, and collaboration with the full teaching staff. As an example, School B students may have felt more involved in important decision making if this was done, resolving the significant variance in Survey #2’s Item #7- Opinion Before Decisions.
Another finding was that students in older years struggle in certain aspects of tutorial inclusion. This is aligned with the literature, and in particular, with Salzman et al. (2019). The researchers found in a social network analysis study of an undergraduate course that year predicted connectedness in the classroom network. This is aligned with this study’s survey findings whereby Year 3 students struggled to feel connected to peers in Survey #1 and Year 2 students felt less able to influence tutorial decisions and had the lowest levels of community in Survey #2. Related to this, students who were involved in university activities struggled with peer interactions and feeling that they could contribute their opinion. It is possible that these students may have felt less included in the tutorial groups simply because of their assumed higher levels of inclusion outside of the boundaries of the classroom at university. In future work, it would be interesting to see how university involvement relates to commuter students too. For these students, it is possible that they do not have time to be involved due to living further away. Therefore, participation in the classroom is their university involvement, rather than societies, sports teams, or other campus extracurriculars.

This research also found that gender was never significantly related to lower average MBIE scores. The literature emphasises that male students generally will participate more in class discussions, while women find their gender identity to influence their course-related self-efficacy and how teachers interpret their behaviour (Aguillon et al., 2020). Based on this, I expected gender to be equally salient in students’ perceptions of inclusion, conceptualised as participation and contribution. This study’s results that counter the literature may have to do with the fact that the classrooms, programmes, and university as a whole had a female majority. Opie et al. (2019) found that female students when they’re in the minority will also achieve higher participation grades when they have female teachers. This highlights how the fact that female students were the majority and that the majority of tutors were female as well may influence the lack of gender salience in the MBIE results. To fully test whether gender dynamics are unrelated to perceptions of inclusion at the university, the study would need to be extended to a school or tutorial groups at the university that are more male-dominated. It is interesting though that this study showed how gender remains salient when it comes to how much importance students place on safety and community within their tutorial groups. For importance placed on community (scale of 1-5), females scored on average 3.91, while males reported 3.87 and non-binary/trans students a lowly 2.75 (p=0.0385). A similar pattern was seen for importance placed on safety (scale of 1-5) with females averaging 4.42, males averaging 4.23, and non-binary/trans students averaging 3.50 (p=0.0352). This hints at perhaps that males and non-binary/trans students feel safety and community is more important in other domains of their personal life beyond that of the tutorial. In this way, while the results indicated gender is not relevant in perceptions of inclusion, it would still be worthwhile to
understand how gender and sentiments of inclusion/exclusion extend to other facets of the student experience whether that be other classrooms, schools, student accommodations, or student societies and extracurriculars.

Interestingly, non-heterosexual students and students with disabilities showed significantly higher levels of inclusion on certain MBIE statements. This was mainly in reference to Item #6 - Last to Know. On this survey item, in Survey #2, both groups shared that they feel they are not the last to know about important changes in the tutorials. Rather, heterosexual students and students without disabilities felt they were the last to know. This is counter to the literature which details that students with disabilities and students from sexual minorities experience lower levels of inclusion. Ferfolja et al. (2020) found in one study of 1,980 Australian sexual minority students, 19% of students experienced exclusion within the last calendar year, and 28% of those fell victim to repeat exclusion. Similarly, Osborne (2019) found in their study of 136 university students with disabilities that 16% felt unwelcome and excluded in classroom activities. My opposing results could be for a few reasons. First, there is a chance that tutors after their inclusion trainings were hyperaware of students who were part of minority groups. This may mean that they over-catered to minority groups in their communications. This could especially be true for students with disabilities as tutors receive a learning adjustment list where each student and their disability are made known to the tutor. As a result, there is a chance that tutors could be displaying a bias towards students in sexual minority groups and students with disabilities. More research would need to be done to confirm this though.

The final major finding in this study was in terms of racial and ethnic minorities’ experiences of inclusion, whereby students’ racial and ethnic demographic group proved to be the most salient demographic in inclusion/exclusion scores on an item-by-item basis. This finding anticipates the major finding of Chapter #7. Before discussing this though, it is first important to understand how the literature paints a picture of international students, and namely Chinese students, as entering UK classrooms only to realise they are hostile environments (Lomer & Mittelmeier, 2021). This then causes these students, who make up the majority of the BAME students in this study, to shut down (Lomer & Mittelmeier, 2021). It is critical to also keep in mind the lasting impact of COVID-19 and Chinese university students’ resulting experiences of xenophobia (Al-Talib et al., 2023). Even as the literature often paints this exclusionary picture of Chinese students, my findings align with the literature that being racially and/or ethnically “different is not deficient” (Lomer & Mittelmeier, 2021, 2). This study shows that there was no significant difference in terms of MBIE cumulative scores for students identifying as Asian in comparison to white students. While there were some aspects of the survey that showed Asian students do feel the classroom could be more inclusive, the overall difference in perceptions of inclusion was negligible. It was clear at the same
time though that Asian students do not feel as invited to contribute their opinion by tutors as white students or BAME and non-Asian students. In a study of BAME students who dropped out of university, Kauser et al. (2021) found that many students reported a feeling of being overlooked by staff, including tutors. Furthermore, white students at times felt disconnected and excluded, especially when it came to communication about meetings with course organisers. Tutors, to an extent, were aware of their lack of these types of communications though. This shows that perhaps communications like this need to be explicitly made by both course organisers and tutors, not just the former. Otherwise, it comes across as disjointed and exclusionary.

6.4: Limitations

A few limitations could impact the conclusions drawn in this chapter. To begin, students’ expectations for inclusion in tutorials could shift over time. This is especially true for first years, who may come into university with a certain idea of what tutorials are and what the dynamics of tutorials will be like. As the term progresses, first years could adapt their expectations as they settle into the new atmosphere of university. This could impact how they perceive inclusion in the classroom. Another minor limitation for this study is that participation did significantly drop over time. Survey #2 had a participant completion rate of 38.04% for School A. This was likely directly linked to the drop in student attendance over time. In the last data collection week for both schools, there were 79 student absences and 61 students who were impacted by the strike. This left just 113 students who attended their regularly scheduled tutorials, which probably impacts who filled out the final survey. If this research study were to be repeated, it would be ideal to run focus groups or another form of interview with students to understand what their expectations were for tutorials and how strikes impacted their perceptions of inclusion. This might help alleviate any chance where students who are feeling especially excluded or altogether are skipping class may not be filling out the survey at all.

Methodologically, this study could be improved if repeated in the future. It would be ideal to incorporate a control group in order to relate the findings back to the specific inclusion training tutors were given. School could be used as a proxy to an extent in the findings. As an example, in Survey #2, School B students had lower average scores for Item #7- Opinion Before Decisions. This could be linked to the training itself among other reasons, but it is hard to tell without a control group and considering how heavily strikes impacted School B’s students. Related to this, this study did not run any focus groups or interviews with tutors themselves due to constraints on the researcher’s hours in conducting two studies in the same term simultaneously. It would be useful to
understand how tutors feel that the training prepared them for fostering inclusion within their tutorial groups.

If the survey portion of this study were to be repeated, it could be insightful to ask a few more questions of students. This would help explore if there are other factors at play in terms of students’ perceptions of inclusion. While the study asked about students’ race and ethnicity, it may also be advantageous to ask about students’ international or national student status and linguistic ability. This is inferred from the data, but not known for certain. Similarly, other demographic questions could include whether students are classified as commuter students, have caretaking responsibilities, and are balancing a job(s) and coursework. Part of the reason UCU strikes occurred was due to the rising cost of living in the UK. This of course may impact students and their decision to pursue both university and a job at the same time. Lastly, while it is known how many tutorial sessions were cancelled due to the strike in these two courses, it is unknown how many of students’ other courses were impacted. All of these factors could play a role in students’ perceptions of inclusion within their specific tutorial groups and their relationships with their tutors.

6.5: Chapter Conclusion

This chapter uses inclusion survey data and uniquely is the first study to apply an adapted version of the Mor Barak Inclusion-Exclusion Scale to higher education. In doing so, this chapter seeks to understand student perceptions of inclusion, conceptualised as participation and contribution. By surveying large groups of students from two schools’ foundational courses, this study is able to investigate group-level and covert social dynamics. Statistical analysis was used to understand potential relationships between demographic groups, composite MBIE scores, and individual MBIE survey items. The demographic groups analysed included gender, race and ethnicity, sexuality, disability status, university involvement, year in school, and school. No significant relationship between demographic group and MBIE composite scores was found for either survey, alluding to a lack of evidence for any grand scale inequality regime being upheld in the classroom. However, using item-by-item statistical analysis showed that race and ethnicity is the most salient dimension of diversity within the classroom when it comes to inclusion and exclusion. Even though composite MBIE scores were significantly higher in Survey #2 than in Survey #1, this chapter concludes with a discussion of how tutors may change their behaviours in order to more effectively foster inclusion in the classroom for all students regardless of identity.
Chapter 7: Quantifying Inclusion with Social Network Analysis

As discussed in Chapter 1, inequality regimes are when organisations adopt and reconstruct existing social stratifications (Acker, 2006). This chapter is concerned with how individual classrooms adopt and adjust the hierarchical status quo as part of a higher education institution. Individual classrooms, in this case tutorials, may be seen as working groups within the higher education organisation. Within the three working groups studied, this chapter focuses on inequalities between racial and ethnic demographic groups. While other demographics are of course relevant in inequality regimes, Chapter 6 highlighted how race and ethnicity were the most salient dimension of diversity for differences in perceptions of inclusion. Moreover, equipped with the skills and tricks to promote inclusion, tutors were in a unique position to create educational spaces where exclusion and thus inequalities between racial groups could be mitigated.

Inclusion, defined in Chapter 2, is both a person’s uniqueness and belongingness, as well as their participation and contribution (Roberson, 2006; Shore et al., 2011). In applying Van Dijk & Khattab’s (2021) legitimacy theory of inclusion, this chapter is concerned with the objective and subjective experience of inclusion/exclusion. Applying the new theoretical framework, I strive to look at the group-level and covert manifestation of inclusion/exclusion. To comprehensively measure inclusion/exclusion then, this study asked participants to take part in the Mor Barak Inclusion-Exclusion (MBIE) Scale survey, and also recorded tutorial sessions. This chapter focuses on the latter to investigate covert group-level behaviours that create inclusion/exclusion, as it relates to participation and contribution. Looking at more covert behaviours such as speaking times and interruptions, I uncover whether participation and contribution occurs along racial lines and discuss how researchers may need to alter the way we conceptualise inclusion.

The larger context of race and ethnicity within the university classroom is important to understand potential lines of friction. For 2021/2022, 29.16% of the university’s entrants identified as BAME. This chapter also considers race and ethnicity, to a certain extent, as a pseudo-proxy for international student status because of the large overlap of Asian ethnicities with international student status at the university. This study, unfortunately, did not collect any data on whether students identified as international students or not, and this is a study limitation. From public records though, the university reports that 31% of all students are international and consequently pay international student fees. Of these, the largest national demographic is Chinese. The other largest non-European international student groups are from the U.S., Malaysia, Hong Kong, Canada, and India.
Entering data analysis, I anticipated an inequality regime of international students of Asian descent being excluded in tutorials given my personal observations, findings from Chapter 6, and the documentation of these patterns of exclusion in the literature. In fact, there has been extensive research on why Chinese students in particular may fall into a pattern of exclusion in the UK classroom. The literature finds that international students, especially those from China, do not orally participate in the classroom largely because of language barriers (Heng, 2018). A startling study found that international students in an Australian university felt that tutors preferred native English speakers (Marlina, 2009). Other confounding explanations offered by students themselves include because they feel ignored, they need more time to digest discussions, they need more time to form responses, because they’re nervous, and because of an unfamiliarity with the argumentative norm of Western discourse (Heng, 2018; Marlina, 2009; Straker, 2016). Not to be forgotten, the impacts of the global COVID-19 pandemic also left Chinese international students in the UK feeling more isolated and at higher risk of anti-Asian racism (Liu et al., 2022).

While studying participation and contribution by analysing both demographic and covert dialogic patterns is rare, this chapter employs social network analysis to do just that (Wagner & González-Howard, 2018; Zhang et al., 2022). I specifically outline potential social network analysis metrics that can be used in discussion-based contexts to offer an underutilised lens through which classroom dynamics and covert hierarchies may be investigated. With that, interaction networks are co-created by students and tutors through dialogue in the classroom. The networks visually map who is speaking to each other and who is not engaging in verbal communication. Such social network analysis metrics I will investigate include who is falling at the periphery of the network, and who is generally finding themselves at the centre. As discussed above, I will also add context with other covert behaviours of inclusion/exclusion beyond speech. This includes student and tutor speaking times and interruptions, but also tutors’ use of students names and any touching that may have occurred by tutors. This offers critical information for educators to understand whether behaviours like this are contributing to patterns of inclusion/exclusion within the classroom. I will also continue to bring in the subjective experiences of students into the network visualisations, adding to Chapter 6, by combining both behaviours and perceptions of inclusion/exclusion.

To wrap up this chapter, I recap this chapter’s main findings in terms of network analysis, centrality measures, speaking times, and interruptions, as well as the theoretical implications for inclusion researchers and our conceptualisations of inclusion/exclusion. I therefore showcase how a social network analysis methodology can illuminate group-level and covert behavioural manifestations of inclusion/exclusion. I also provide an ample discussion on what the implications of the findings are for tutors and their professional development. Of course, there are also implications for
higher education institutions as a whole. In particular, I discuss how the undertones of racial inequality in the findings bring up issues with internationalisation at higher education institutions within the United Kingdom. While the chapter is just one study of three tutorials in one higher education institution, it offers insight into potential inequality regimes in place that should be considered concerning by educators.

7.1: Data

The data for this chapter were collected during the autumn semester of the 2022-2023 academic year at the partner university. The three 50-minute classes are referred to as Tutorial A, Tutorial B, and Tutorial C and were recorded three times during the term, totalling nine individual recordings. All three tutors were female, with an average previous tutoring experience of three academic terms. All three tutors attended an inclusion training prior to the term’s commencement, but the type of training depended on their course. Each tutorial group had between 11-14 students on the tutorial's student roster. These numbers fluctuated at the beginning of the term as students were added to the course, removed from the course, and requested tutorial switches due to scheduling issues. Attendance, seen in Table 7.1, also fluctuated in the tutorials week-by-week as attendance had no effect on the student’s marks. 41 total participants were involved in this study, of which three were the tutors. Each participant was assigned a unique participant ID #, which combined their tutorial code (ex. A) with a number. All tutors were assigned #1 to ensure it was clear in visualisations who the tutor was.

<table>
<thead>
<tr>
<th>Tutorial Group</th>
<th>Start of Term Recording (# of Students)</th>
<th>Mid-Term Recording (# of Students)</th>
<th>End of Term Recording (# of Students)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tutorial A</td>
<td>10</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>Tutorial B</td>
<td>9</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>Tutorial C</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

As mentioned in Chapter 3, recordings were collected with an iPad for each individual tutorial session. The nine recordings totalled 7 hours, 21 minutes, and 10 seconds of video footage. It should be noted that recordings usually began prior to and
ended after official class times. This is due to me as the researcher having to come in and set up prior to class time, as well as ending the recording after class officially ended. One recording was cut short for Tutorial B due to a technological malfunction, with an estimated 10-15 minutes left of class time. Furthermore, small group activities, while recorded, were not transcribed nor included in the following analysis as students were speaking over each other. With only one recording device, it was impossible to discern who was saying what. It would be extremely worthwhile to study these smaller group encounters which Diehl & McFarland (2012, p. 336) find encourage “opportunities for identity validation”. This research is more preoccupied with the full class discussions when everyone could respond to everyone, and when speaking times could be more accurately discerned.

Transcriptions were initially carried out by Otter.ai and I then personally went through each transcription two more times. Given Otter.ai is an American platform and sometimes struggles with less American accents, manual transcription was of the utmost importance to ensure accuracy. Speech was unintelligible if a speaker was speaking too softly, mumbling while wearing a mask, or if multiple speakers were speaking over each other. Total speaking times were unaffected by this, as I could still determine who was speaking, and for how long they spoke. To determine who was speaking to who, I relied on the video recordings. Using each other’s names was rare, so determining who was speaking to who was largely determined by a mixture of eye contact, body language, body orientation, dialogue (ie. use of ‘you’, etc.), and the context of the dialogue.

Some survey data are used in this chapter that was also used in Chapter 6. This includes the Mor Barak Inclusion-Exclusion Scale scores and demographic data. As mentioned in Chapter 6, the MBIE Scale is grounded in perceptions of participation and contribution. Questions such as “I am invited to actively participate in communication with my tutor” were asked to students, while the questions asked to tutors were re-framed to the tutor experience. For a comprehensive understanding of how the scale was adopted and administered to participants, please see Chapter 6. Unfortunately, some students never completed a survey and did not respond to additional follow-up emails after the term to complete demographic surveys. This means that of the three tutorials studied in this chapter, four students all from Tutorial A never submitted demographic information. However, Tutorial A is the class I personally sat in. As a result, in this chapter, I categorise the four remaining students in terms of race. My categorisation is based on personally meeting and talking with the students, knowing their names, conversations with the tutors, and hearing them speak certain languages. Although I am fairly confident in my categorisation of students, there is always potential for miscategorisation. Consequently, I have also included the network visualisations without these classifications in Appendix C.6. The overall breakdown of students was
48% BAME students (predominantly Asian, and a couple Middle Eastern and Latinx students) and 52% white students. This compares to the university reporting 64% white identifying entrants in 2021-2022, meaning BAME students may be overrepresented in this study. Lastly, two tutors were white and one identified as Asian.

Figure 7.2: Racial Demographics of Tutorials (Not Including Tutors)

7.2: Social Network Analysis

Social network analysis is used to visualise the speaking relationships within each of the three tutorial groups. A ‘node’ in this case, is one individual participant. Uniquely in classroom settings, there are also cases whereby the tutor will speak to the whole class and not one individual student. In this case, there is an additional ‘Class’ node added to the dataset. The addition of this node allows analysis to distinguish between student-student and student-teacher interactions, rather than creating edges to each node that may overshadow these relationships (Mameli et al., 2015). An ‘edge’ is the speaking that is done from one person to another person or the full class. To code the speaking, or edge, I relied on the video recordings for signals of when a conversational turn was over. Edges demonstrate two other features of the relationship between people: directionality and frequency (Wagner & González-Howard, 2018). Directionality showcases who the speaker is and who the receiver is in this dataset (Wagner & González-Howard, 2018). Frequency represents the total number of edges between two participants, whereby a higher frequency indicates more communication (Wagner & González-Howard, 2018).

Together, nodes and edges are pulled together into an adjacency matrix which provides the foundation for network visualisations and further statistical analysis (Scott, 2011). As demonstrated in the adjacency matrix in Table 7.3, the tutor (C1) spoke to the full class (Class) 37 times. The full class never provided a “choral response” as Diehl &
McFarland (2012, p. 333) refer to, which is to be expected given cases where the full class speaks in unison to the facilitator is more common in elementary education. This only happened once across the three tutorial groups. While the adjacency matrix below is directional, adjacency matrices may also be made that are non-directional. In that case, the sums of all interactions between two individual nodes are summed. It feels important to distinguish directions in a classroom, where certain students may be overwhelmingly carrying the discussion. To create the nine adjacency matrices, I used the R programming language and the igraph package (Csárdi, 2023). I also used the packages tidyverse (Wickham, 2016), netUtils (Schoch, 2013), and tnet (Opsahl, 2020).

Table 7.3: Example of an Adjacency Matrix (Tutorial C- Recording #2)

<table>
<thead>
<tr>
<th>NODE</th>
<th>C5</th>
<th>C10</th>
<th>C1</th>
<th>C8</th>
<th>C12</th>
<th>Class</th>
<th>C4</th>
<th>C2</th>
<th>C9</th>
<th>C11</th>
</tr>
</thead>
<tbody>
<tr>
<td>C5</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C10</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>37</td>
<td>43</td>
<td>37</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>C8</td>
<td>0</td>
<td>0</td>
<td>45</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C12</td>
<td>0</td>
<td>0</td>
<td>58</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Class</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C4</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C2</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C9</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Certain node and network properties are relevant in understanding how a network may look if it was perfectly inclusive. Figure 7.4 demonstrates a hypothetical inclusive and exclusive network with five nodes, for simplicity's sake. In the inclusive network, each node is connected to every other node meaning the overall network is decentralised with no core or periphery. Conversely, there are no isolates, which means no nodes exist that have no connections whatsoever. In the classroom setting, decentralisation would translate to each student speaking to every other student. In other words, Student A may say something to Student B, and Student B also speaks to Student A at some point too. In a university classroom with 10+ students though, this becomes more difficult to achieve. Each edge is also of equal width, indicating an equal frequency of communication between each student. In the aforementioned example, Student A may speak to Student B five times, but Student B also speaks to Student A five times. No singular node is dominating the conversation. This relates to the concept of degree centrality, which demonstrates how active or popular a node is within the network by showing how many relationships each individual has (Landherr et al., 2010). In a perfectly inclusive network, there would be no variance in degree distribution across
the network. In a directed network like the ones in this chapter, degree centrality is broken up into in-degree and out-degree, which represents dialogue being received and dialogue being sent, respectively. In that way, in-degree can be seen as representing popularity, while out-degree may be seen as representing activity.

To an extent, a perfectly reciprocal and decentralised network in the classroom cannot be expected. For instance, imagine a typical tutorial scenario where a tutor asks a question to the class. Student B never responds to questions, but Student A responds. The tutor says “Great answer”. It is unlikely Student A will then respond to that just to ensure perfect reciprocation. This would automatically translate into a centralised network with a core and periphery network, as well as unequal in-degree and out-degree across the network. This may be seen in the exclusive network of Figure 7.4, where node H4 speaks to H2 and H3, but H2 and H3 never reciprocate. Other exclusive attributes of the network are the heavier weighted edge between H2 and H3, indicating more frequent dialogue between these two nodes. H1 is an isolate node, meaning H1 never spoke to anyone and was never spoken to. A scenario where this may arise is if a student’s English is less strong, meaning they hesitate to speak up and are never spoken to. This puts forth just one example of why it is important to understand how dimensions of diversity relate to one’s inclusion within a network. Understanding if there are racial and ethnic patterns to where individuals fall within these inherently unequal networks is the purpose of this chapter.

Figure 7.4: Example of an Inclusive Network

While deciding whether a network will be directional or non-directional is a crucial first step in creating a network, it is also important to decide whether to weight or not
weight the network as well. A non-weighted network will not combine the individual edges between nodes. A weighted network will collapse all edges of the same direction into one edge such as in the inclusive network in Figure 7.4, weighting the thickness of the edge by the number of edges there originally were. Weighting the network made the most sense in the case of the classroom for the same reason direction felt important. Simply, it felt important to know whether conversations were being reciprocated, even if not equally as discussed above. If a tutor speaks to a student, is that student at least responding? Alternatively, if a student is dominating conversation, are others engaging back? Furthermore, the weighted values were normalised by dividing the weights by the maximum weight in the network, then multiplied by 5. This allowed for edges to be visually distinct by weights in network visualisations.

After weighting the network, two attributes were layered onto the nodes for further analysis: the Mor-Barak Inclusion-Exclusion Scale scores and the race/ethnicity of the participants. The MBIE scores were taken twice during the study, during the same week as the first tutorial recording and the same week as the last tutorial recording. Race and ethnicities of participants were only obtained once during the same week as the first tutorial recording. If students did not fill out the first survey, they were asked later in the academic year to fill in a survey solely regarding their demographics. While students reported their race and ethnicities explicitly from an extensive list, the following encompassing and categorical groups were used to assign nodes: No Race/Ethnicity Reported, White, Latinx, Asian, and Middle Eastern. These groups are large enough to encapsulate the participants’ identities while protecting their anonymity, but of course, do not accurately represent participants’ intersectional identities. While there were Black participants in the larger study, there were no Black participants in the three tutorials that were recorded. An example of one of the final individual networks is below, as well as a combined network. The combined networks do not reflect an MBIE score, but combine all the data from the three recordings for one tutorial into one network. For the full overview of the composite network and each individual network for the tutorials, please see Figure 7.5 through Figure 7.10, and Appendix C.6 for mid-term networks specifically.

7.2.1: Core-Periphery Analysis
Figure 7.5: Cumulative Network for Tutorial A

Figure 7.6: Cumulative Network for Tutorial B

Figure 7.7: Cumulative Network for Tutorial C
When all three tutorial sessions are condensed into the one composite network per Tutorial A, B, and C, two patterns emerge at first glance. First, the tutor is always at the core of the network. This is to be expected given tutors act as facilitators of class discussions. It is notable that Tutorial B saw the most amount of student-to-student conversations, showing that the tutor was able to encourage student-to-student dialogue in the full class discussions. Moreover, the second pattern that can be seen is that Asian students do tend to fall towards the periphery of the networks. To illustrate, in the composite network for Tutorial A, it is clear that most two-way conversations occur between the white students and the white tutor (namely, A2 and A3 converse back and forth with A1, the tutor). The same can be seen in Tutorial C where most conversation occurs between C1 (the tutor), C8, C12, and C3. C9 is the only Asian student who could potentially be considered as part of the core in the composite network for Tutorial C. Notably, Tutorial C is the only tutorial with a BAME tutor, but the core being majority white students remains. The only isolate in the composite network is also an Asian student, C11. Interestingly, there was more racial and ethnic diversity in Tutorial B, with white, Asian, Latinx, and Middle Eastern students. Unlike Tutorials A and C, there seemed to be less clustering around the tutor, B1, that seemed to occur based on racial/ethnic identity. Even so, all three isolates (B12, B13, and B14) in Tutorial B are
Asian identifying. This means of the four isolates across all three tutorials, all four were Asian students.

To confirm that tutors and white students dominate the core, I ran a core-periphery test on all three composite networks. The core-periphery test identifies which nodes belong to the core of a network, and which likely belong to the periphery of a network. I ran this test twice, once weighting by degree centrality and once weighting by Eigenvector centrality. Eigenvector centrality shows which nodes have many relationships as well as if those relationships are with other nodes who themselves maintain many relationships (Landherr et al., 2010). It can therefore be likened to who has enough social capital in the classroom to belong to the core. Both core-periphery tests confirmed that the tutors (A1, B1, and C1) are part of the core, which is to be expected given they are the facilitators of class discussions. The degree centrality driven core-periphery test identified that 66.67% of the core identifies as white, while 33.33% identifies as BAME. The periphery, on the other hand, was made up of 27.27% white students and 72.72% BAME majority Asian students. According to an Eigenvector centrality core-periphery test, the only BAME node who was part of the core was C1, the tutor from Tutorial C. These findings emphasise a disparity within the classroom whereby even though 56.76% of the class, including tutors, was BAME identifying, the core was majority white identifying.

Context-wise, it is also crucial to comment on how the physical rooms and topics of conversation differed for the tutorials. For Tutorial A, there was one long table made up of small tables and students sometimes moved the tables into smaller islands. Tutorial B had a U-shape layout and students never moved the tables. This meant everyone could see everyone else all times in all three sessions. Furthermore, students generally sat in the same seats each session for Tutorial B. Tutorial C had a similar classroom layout as Tutorial B in that students sat in generally the same seats each session in a U-shape formation. There may be a possibility that the physicality of the room helped fuel classroom dynamics, where Tutorial B seems to show students engaging more with each other. Related to this, it is also seen in Tutorials A and C that some Asian students only ever spoke with each other. This is especially clear in Tutorial C where student C10 speaks back and forth with another Asian student C5, but chooses not to speak back to the tutor, C1, when spoken to. This also brings forth the question of how topics of class discussion could influence dynamics. To demonstrate, Tutorials B and C had one tutorial devoted to social class, which students from the UK (majority white students) may have more knowledge of to begin with. Tutorial A, on the other hand, had a class devoted to the topic of global enterprises and greenwashing, which could be seen as a more inclusive topic of conversation.

The above discussion brings into question whether MBIE scores are impacted by the core and periphery split, given Chapter 6 found race/ethnicity as the most salient
dimension of diversity when it comes to MBIE scores. Yet, using a Welch two-sample t-test, core/periphery identification is not significantly correlated to students' MBIE scores for the start (p= 0.7466) nor the end of term (p= 0.3119). Looking at the individual tutorial networks highlights this as students who are at the periphery of the network do not always experience lower levels of inclusion. Conversely, just because a student is at the centre of discussion, as many white students were, does not mean they automatically experience higher levels of perceived inclusion. For example, as seen in Figure 7.8, A3 is one of the closest to the tutor in the start of term network, A3 reports the lowest level of inclusion (3.45). Of the three students (A2, A3, A5) who have the most high-traffic conversations with the tutor in that session, not one of them has the highest MBIE score. Rather, A11 who never speaks to any student nor the tutor has the highest inclusion score (4.81). A11 also has the highest inclusion score by the end of term session, which increases to a 6.0. A11 is seen to begin speaking to the tutor and other students by that last session too. To add context to that last session, the session’s group was noticeably smaller with fewer students in attendance, and the nature of the activity that day was working together in a full-class small group.

This pattern of central students not always experiencing high sentiments of inclusion is also noticeable to a certain extent when comparing Tutorial B and Tutorial C. While Tutorial B’s tutor has less immediate clustering correlated with racial/ethnic identity, it is noticeable that the lowest two MBIE scores in the tutorial by end of term are both Asian students. Tutorial B also saw students who remained isolated throughout the term. Namely, B13 never speaks in group discussions and is never spoken to. Unsurprisingly, B13’s MBIE score remains quite low by the end of term, jumping a marginal +0.19 from 2.36 to 2.55. Perhaps, this is due to the fact that attendance never drops off as it did in Tutorial A, keeping the isolated students on the margins. Yet, in direct contrast to Tutorial B, C10 and C11 isolates in Tutorial C have some of the highest inclusion scores by the end of term. The two lowest inclusion scores are C4 and C2, both with 3.55 on the MBIE scale. C2 and C4 both engage in back and forth conversation with the tutor. C2’s MBIE score actually decreases over time despite the consistent engagement. Accordingly, it is once again demonstrated that more contribution and engagement does not automatically translate into higher levels of perceived inclusion. All of this taken together prompts whether there is something tutors can do to engage students with low levels of inclusion both at the periphery and at the centre of class discussions.
Figure 7.8: Start and End of Term for Tutorial A with MBIE Scores

Figure 7.9: Start and End of Term for Tutorial B with MBIE Scores
7.2.2: Centrality Measurements

Centrality measurements offer insight into who the key individuals within a given network are (Landherr et al., 2010). In the classroom, centrality measurements offer insight into to what extent the tutor is a key individual, as well as if other students are acting as central figures. There are a wide array of centrality measurements though. The three most important to this study are degree centrality, closeness centrality, and betweenness centrality. Closeness centrality is how proximal a node is to other nodes in a network (Landherr et al., 2010). The tutor would most likely have the highest closeness centrality in a tutorial as the course facilitator. The highest betweenness centrality would also likely represent the tutor, whereby the node with the highest betweenness centrality maintains the shortest path between all other nodes (Landherr et al., 2010). The node with the highest betweenness centrality can therefore be seen as a controller, controlling the flow of resources, information, and communication.

I used R to calculate each of these centrality measures. In-degree, out-degree, closeness centrality, and betweenness centrality were calculated with the igraph package (Csárdi, 2023). Closeness centrality for Tutorials B and C were calculated utilising the tnet package (Opsahl, 2022). This is because igraph’s closeness function assumes that the network is well connected. As seen above, Tutorials B and C both had isolates, and as a result, they were not well connected networks. The tnet package offers a function that takes this into account. Furthermore, some calculations were
based on an unweighted network except for the closeness calculations for Tutorials B and C. While 'Class' was kept in the calculations for the in-degree measurement, it was ignored as a node in the calculations for out-degree, closeness centrality, and betweenness centrality. Table 7.11 through Table 7.13 rank the top five nodes in each tutorial per centrality measure, while highlighting other contextual factors such as a tutor mentioning a student by name or a student being patted on the back by a tutor.

### Table 7.11: Centrality Measurements for Tutorial A

<table>
<thead>
<tr>
<th>Top Five</th>
<th>In-Degree (non-weighted, directional)</th>
<th>Out-Degree (non-weighted, directional)</th>
<th>Closeness (non-weighted, directional)</th>
<th>Betweenness (non-weighted, directional)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.</td>
<td>A1: 12.14</td>
<td>A2: 13.93*</td>
<td>A2*/3*: 0.048</td>
<td>A2*: 0.02882</td>
</tr>
<tr>
<td>3.</td>
<td>A2: 9.00**</td>
<td>A3: 11.71*</td>
<td>A14/10: 0.043</td>
<td>A3*: 0.01021</td>
</tr>
<tr>
<td>4.</td>
<td>A3: 6.29*</td>
<td>A7: 1.71*</td>
<td>A7*/11: 0.042</td>
<td>A11: 0.00009</td>
</tr>
<tr>
<td>5.</td>
<td>A11: 1.38</td>
<td>A11: 1.43</td>
<td>A4/5/9/6/8/12/13*: 0.040</td>
<td>N/A</td>
</tr>
</tbody>
</table>

^ Tutor patted student on back  * Name directly mentioned by tutor

### Table 7.12: Centrality Measurements for Tutorial B

<table>
<thead>
<tr>
<th>Top Five</th>
<th>In-Degree (non-weighted, directional)</th>
<th>Out-Degree (non-weighted, directional)</th>
<th>Closeness (weighted, directional)</th>
<th>Betweenness (non-weighted, directional)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.</td>
<td>Class: 7.00</td>
<td>B9*: 2.42</td>
<td>B9*: 8.00</td>
<td>B9*: 0.00061</td>
</tr>
<tr>
<td>3.</td>
<td>B8: 2.17</td>
<td>B8: 2.08</td>
<td>B8: 7.53</td>
<td>B8: 0.00040</td>
</tr>
<tr>
<td>4.</td>
<td>B9*: 1.84</td>
<td>B4: 1.92</td>
<td>B4: 7.30</td>
<td>B5: 0.00009</td>
</tr>
</tbody>
</table>

* Name directly mentioned by tutor
Looking at Table 7.11 through Table 7.13, the centrality measures confirm a few patterns that to a certain extent can be seen in the network itself. Across the board, it is verified that tutors occupy the most central position in the network. Tutors rank #1 in closeness centrality showing that they maintain the closest position to all other nodes. However, in Tutorial A, the highest in-degree is actually held by the ‘Class’ node. This could potentially be driven by the tutor addressing the full class more, instead of responding to or speaking directly to individual students. In a discussion, this may signify more of a lecture-based discussion. For betweenness centrality, it was also confirmed that the tutor is indeed the controller in tutorials, controlling the information and resource flows. Another pattern potentially missed in the network visualisations is who is driving discussions on a whole. In terms of out-degree, it is seen that Tutorial B maintains less of a disparity between the amounts which students are contributing. In Table 7.11, it is seen that A2 and A3 drive much of the student participation in Tutorial A, while Table 7.13 shows that C12 is the equivalent driver in Tutorial C. Yet, in Tutorial B, there are less stark differences between the top ranking students in terms of out-degree. This may signify that tutor B1 was the most successful at encouraging more students to participate equally.

After identifying who the most central figures were according to the different centrality measures, it was next important to investigate whether there were underlying patterns, especially in terms of students’ racial identities. It is important to note that the analyses I completed to do this did not control for tie dependency. Even so, I ran a linear multiple regression analysis to understand whether the centrality figures had any significant relationship with students’ MBIE scores. There was no significant relationship between any of the four centrality measures and MBIE scores at the start nor end of term. This once again emphasises how more contribution and engagement, or inclusion behaviours, do not automatically translate into higher levels of perceived inclusion. With a Mann-Whitney Wilcoxon test, I also checked whether the centrality
measures have a significant relationship with students’ racial identities, as the core-periphery analysis did. In-degree (p= 0.0051), out-degree (p= 0.0024), and closeness centrality (p= 0.0079) were each significantly related to students’ status as either BAME or white. Betweenness did not prove significantly related. Taking these analyses together indicates that once again that students’ racial identity is significant in understanding how inclusion manifests in terms of behaviours within the classroom. At the same time, it is less salient when it comes to students’ perceptions of inclusion, which Chapter 6 explored in depth.

Also worthwhile to note is that recordings were coded for any touching and name-calling too. Name-calling entails the tutor addressing a student by name, outside of attendance roll call. When this is overlaid with centrality measures, a few noteworthy patterns emerge. First, tutor B1 used student names the least. In fact, B1 only addressed a student directly by name once to acknowledge a point a student made. Interestingly though, B1 had the most students speaking across the board. This potentially points to the fact that students do not need to be directly called upon by name to feel comfortable regularly speaking. Yet, the one student, B9, the tutor did refer to by name did rank the highest of all students in 3 centrality measures. C1 utilised student names a total of six times; C12 twice, C10 twice, and C5 and C9 once each. Three of these students were Asian, signalling that C1 who also identified as Asian, may feel more comfortable with other Asian names. A1 used student names the most outside of attendance roll call. A1 referred to A7 four times, A13 three times, A3 three times, and A2 once. Three of these students were highly central in class networks and two of these students were Asian. A13, on the other hand, occupied a less central position in class discussions. The context in which A1 used A13’s name was mostly in checking in with the student who arrived late one session, and in an attempt to check in with the student towards the end of the tutorial. C1 is also the only tutor to touch a student, having patted A2, arguably the most central student in the tutorial network, on the back. All together, the centrality measures and the name calling analysis points to the fact that tutors utilising students’ names directly is largely underutilised. If a tutor does refer to a student directly, it is usually someone who is central to tutorial discussions. The lack of name usage by tutors could also point to the fact that tutors have many students, with up to four tutorial groups per term. This could mean any one tutor could have about 60 students, and 60 names, to learn over the course of ten weeks. In future research, name calling analysis may prove worthwhile for further study, especially in international classroom contexts.

7.2.3: Speaking Times
Table 7.14: Top Five Speakers per Tutorial (Minutes)

<table>
<thead>
<tr>
<th>Top Five</th>
<th>Tutorial A</th>
<th>Tutorial B</th>
<th>Tutorial C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>A1: 39.93</td>
<td>B1: 36.6</td>
<td>C1: 62.90*</td>
</tr>
<tr>
<td>2.</td>
<td>A2: 12.77</td>
<td>B8: 7.48</td>
<td>C8: 17.08</td>
</tr>
<tr>
<td>3.</td>
<td>A3: 9.05</td>
<td>B4: 6.00</td>
<td>C12: 5.63</td>
</tr>
<tr>
<td>4.</td>
<td>A4: 3.02*</td>
<td>B9: 5.63</td>
<td>C3: 2.37</td>
</tr>
<tr>
<td>5.</td>
<td>A7: 1.95*</td>
<td>B5: 4.63*</td>
<td>C2: 2.23</td>
</tr>
</tbody>
</table>

* = racial/ethnic minority

Speaking times were also analysed in Table 7.14 to further investigate the patterns seen in the tutorial networks. As to be expected, in all three tutorials, the tutors (A1, B1, and C1) all spoke the most during the class discussions. This is unsurprising given tutors may act as discussion facilitators, asking prompting questions, responding with follow-up questions, and adding relevant course learnings. One study by Hardman (2016) found the average length of a tutor’s turn during a tutorial is 45 words, while students will say an average of 18 words per turn. In this piece of research, students were also found to say 18.9 words per turn. In contrast, the tutors’ average turn was 28.1 words, which is significantly lower than in Hardman’s (2016) study. This may potentially indicate that tutors spoke the most during the class in terms of frequency, but their turns may have been less substantive though. This could perhaps consist more of shorter responses and questions than in the Hardman (2016) study.

Figure 7.15: Total Speaking Times per Racial Demographic Group
What is also noticeable when looking at the top five speakers for each tutorial, as well as Figure 7.15, is the lower speaking times of BAME speakers. In fact, the only BAME student speakers to make it into the top are A4, A7, and B5 for Tutorial A and Tutorial B, respectively. No BAME students in Tutorial C charted in the top five speakers. These findings support what was seen in the core-periphery network analyses as white students tended to be central in the networks. The relationship between total speaking time and racial identification is overall significant ($p= 0.0041$). Other interesting findings when investigating speaking time can be seen in looking at Tutorials B and C. In Tutorial B, there is a much more even distribution of speaking time in the top five speakers. This also reflects the closely knit network in Tutorial B whereby more students engaged in student-to-student discussion. In Tutorial C, the opposite is seen. Instead, tutor C1 dominated full class conversations, nearly doubling the amount of speaking time of tutors A1 and B1. Furthermore, student C8 is seen to have contributed more than the other students. All of this impacts white students’ average speaking time to be 311.5 seconds, while the average speaking time of BAME students was 77.2 seconds over the full observational period.

To add to analysis, and as discussed in this chapter’s data collection section, white students made up 52% of the students in the tutorials studied. Across the three tutorials and not including tutors, 48% of students were BAME. Based off this, it would be expected that similar percentages would emerge in terms of who is contributing and participating in the classroom. On the contrary, white students spoke 74% of the time and BAME students 26% of the time. Furthermore, these breakdowns contribute to the same pattern seen in the social network analysis: BAME and predominantly Asian students on a whole are contributing and participating less in tutorials despite being nearly equally present in class as white students.

Figure 7.16: Breakdown of Speaking Time per Racial Group (No Tutors)
Interruptions also prove to be an interesting point of analysis for the tutorials. Interruptions were coded as when someone would try to take the discussion floor while someone else was still speaking. Tutors were included as perpetrators of interruptions too. Interruptions were not if someone just said “yeah”, “mhm”, or another phrase of agreement or “alliance” while someone was speaking (Fuhse, 2023, p. 1207). Rather, an interruption was seen as a direct attempt to take over as the primary speaker, consequently creating a form of “adversarial” interruption (Fuhse, 2023, p. 1223). With that, exactly 66.67% of the top five speakers per tutorial group interrupted. This may be seen in Table 7.17. Of the study participants who were not in the top five speakers in their respective tutorial group, only 18.18% interrupted. As seen in Table 7.17, Tutorials A and C had more interrupters than Tutorial B. This indicates that perhaps Tutorial B demonstrated a more peaceful turn-taking process than Tutorials A and C. It is also noteworthy that the majority of interruptions were actually only carried out by four people in particular: A1, A2, A3, and C1 with 31, 24, 20, and 9 interruptions each. This hints at perhaps a power struggle in Tutorials A and C. Given the high proportion of interrupters falling into the top five speakers per tutorial overall though, a few questions arise. Primarily, is interrupting a prerequisite for high levels of participation in tutorials? Or, do people who participate just have more time to administer interruptions?

Table 7.17: Top Five Speakers per Tutorial (Minutes) with Interruptions

<table>
<thead>
<tr>
<th>Top Five</th>
<th>Tutorial A</th>
<th>Tutorial B</th>
<th>Tutorial C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>A1: 39.93**</td>
<td>B1: 36.6*</td>
<td>C1: 62.90**</td>
</tr>
<tr>
<td>2.</td>
<td>A2: 12.77**</td>
<td>B8: 7.48</td>
<td>C8: 17.08*</td>
</tr>
<tr>
<td>3.</td>
<td>A3: 9.05**</td>
<td>B4: 6.00</td>
<td>C12: 5.63*</td>
</tr>
<tr>
<td>4.</td>
<td>A4: 3.02</td>
<td>B9: 5.63</td>
<td>C3: 2.37*</td>
</tr>
<tr>
<td>5.</td>
<td>A7: 1.95*</td>
<td>B5: 4.63*</td>
<td>C2: 2.23</td>
</tr>
</tbody>
</table>

* = interrupted someone at least once  
^ = interrupted someone at least five times

Upon further analysis, it appears that white participants are the overwhelming majority of interrupters, and also the majority of receivers. In a hypothetical and perfectly inclusive discussion, it would be expected that if interruptions occur at all, they would likely be evenly distributed amongst racial demographic groups. This is not the case though. White participants interrupted a grand total of 88 instances, while BAME participants interrupted just 14 times. This is displayed in Figure 7.18. Fortunately, although white participants are the predominant interrupters, BAME participants are not the predominant receivers. Rather, as seen in Figure 7.19, 12.75% of interruptions are
targeted at BAME participants. 55.88% are targeted at white participants and 31.37% are towards no one in particular. An example of this is if the students are in breakout groups and someone interrupts breakout group discussions to redirect the class’s attention to themselves. This could be as innocent as asking a clarifying question during group breakouts, whereby a student may ask a question on behalf of their group. Yet, to do so, they have to interrupt the discussions for the entire class. Thus, when analysing interruptions by racial demographic groups, it seems while white participants are the predominant interrupters, it may just be that white participants had more clarifying questions, or were more prone to interrupting given they were already central in discussions.

Figure 7.18: Interruptions Administered per Racial Demographic Group

![Figure 7.18: Interruptions Administered per Racial Demographic Group](image)

Figure 7.19: Interruptions Received per Racial Demographic Group

![Figure 7.19: Interruptions Received per Racial Demographic Group](image)

7.3: Discussion
This chapter presents an exploration of more covert group-level manifestations of inclusion/exclusion within three introductory course tutorials. In doing so, I ground inclusion in observable behaviours of participation and contribution, rather than solely perceptions. This allowed me to convert dialogic data into data appropriate for social network analysis. My main methodological contribution in this chapter is the outlining of key social network analysis metrics that may prove useful for other research into inclusion/exclusion within discussion-based contexts. Such analyses included core-periphery analysis and centrality measures, such as in-degree, out-degree, closeness, and betweenness centrality. Without this analysis, the subtle patterns of exclusion I discussed would not have emerged nor could they be investigated over time without the network visualisations. I further added nuance to my research by still weaving in perceptions of inclusion, as well as tracking behaviours such as name-calling, touching, and interruptions. Currently, the university only monitors attendance. However, a student coming to class is just the first step in engaging in learning. Just because a student attends class does not directly translate into participation and contribution, nor inclusion in that way.

As previously discussed in this thesis, social network analysis is underutilised in the inclusion literature (Wagner & González-Howard, 2018). This chapter offers a comprehensive blueprint for applying social network analysis to inclusion/exclusion research in a way that does not over-rely solely on participants’ perceptions of this social phenomenon. Given demographics influences one’s perceptions of inclusion within professional networks, it makes sense that researchers should be incorporating both students’ perceptions of and behaviours of inclusion/exclusion (Jung & Welch, 2022). Mameli et al. (2015) does this by utilising centrality measures with classroom dialogic data, but my work takes this a step further by outlining how to incorporate demographic analysis and perceptions of inclusion too. I believe that this methodology offers the most nuance in inclusion research in a way that does not disservice participants by overlooking more covert patterns of exclusion.

First and foremost, in terms of findings, if inclusion is conceptualised as participation and contribution, then there are covert behavioural patterns of exclusion within the higher education tutorials studied. These patterns of exclusion are uncovered through social network analysis and would be wholly missed if solely surveys were utilised. Exclusion is seen in the core-periphery analysis whereby when weighted by degree centrality, 66.67% of the participants in the core were white-identifying despite 56.76% of the class, including tutors, identifying as BAME. For example, in Tutorial A, white students cluster around the white tutor, while Asian students sit on the periphery of Tutorial C despite a similarly identifying tutor. The composite network visualisations further show how Asian students at times only speak with each other (ex. C10 to C5).
Three of four centrality measures were also significantly related to race and ethnicity, including: in-degree (p=0.0051), out-degree (p=0.0024), and closeness centrality (p=0.0079). These findings echo Chapter 6 whereby race and ethnicity also proved salient in influencing students’ perceptions of inclusion. These findings stress which racial demographics are the controllers of the emerging tutorial discourse and show who is controlling the knowledge that is being produced in that way. In addition to this, while understanding the exact influence of the room layouts on the classroom dynamics is beyond the remit of this chapter, it is worth noting here that the three individual tutorials included two different types of room layouts. Future research could potentially investigate how classroom layouts help or hinder these subtle patterns of exclusion along racial lines.

These results present a practical implication for higher education in that there must be a reckoning of patterns of exclusion in tutorials and a need, perhaps, to look at the larger context of internationalisation of UK HEIs. In this analysis, of all student contributions, white students were found to participate 74% of the time, while only being 52% of the tutorials studied. BAME students were the contributors 26% of the time, but represented 48% of the tutorial. Of these BAME students, the majority were Asian. This presents a potential inequality regime in tutorials that is race-based, but also perhaps international student status or nationality-based. A study by Rienties et al. (2013) confirmed the salience of the latter as the second best predictor for initial friendship ties in the classroom was nationality. Reflecting on the students’ choices in seating arrangements highlights this potentially nationality- or race-based pattern. Additionally, it is important to keep in mind that the university’s non-EU international undergraduate population has grown from 18.90% in 2011/2012 to 29.13% in 2021/2022, with Chinese students making up the largest demographic.

Another key finding in this chapter directly impacts the way researchers conceptualise both inclusion and participation. This chapter demonstrated that just because someone may fall in the periphery of a network does not mean they automatically feel excluded. In other words, even if a person is displaying significantly lower levels of participation and contribution, they may still experience high perceptions of inclusion. Examples of this exist in both the individual and composite network visualisations. For example, C11, C10, and A11 all are not high contributors, yet they reported the highest perceptions of inclusion. Therefore, investigating covert patterns of inclusion/exclusion at the group-level reveals that these behavioural patterns at times coincide with lower perceptions of inclusion, but not always. This prompts the question of whether we are limiting the way we conceptualise inclusion, by limiting the way we conceptualise participation.

In reflecting on what I think of when I think of participation in a tutorial, I immediately think of talking, speaking, answering questions, and asking questions. I
know at least some other tutors share in this thinking too due to discussions that emerged during the inclusion training intervention sessions. For instance, I remember one tutor asking a question at the end of the inclusion training session of what to do to encourage Chinese students to speak up and participate. This is a direct reflection of how the way we think about participation is inherently limited. In further thinking about my own conceptualisation of participation, I do remember observing Tutorial A and taking note of differences in participation. For example, two students could be staring at their computer during class. To a tutor, it could look like neither student is paying attention to class discussions. From my end though, sitting behind the students, I could take note that one student could be using the computer to translate a word from the worksheet. I could also see the other student is answering text messages through their computer. In this case, one student could be considered to be actively engaging with the class, while the other is not. Yet, to a tutor or any other facilitator who stands at the front of the class, both students would seem to be not participating.

With that, silence can be both non-participative and participative. While the literature lacks studies of overall racial integration within UK university classrooms, there is ample research on the quieter participation of Chinese international students within UK classrooms. Silence as participative is well documented in the literature in that students could be actively listening as participating, but not contributing orally (Elliott & Reynolds, 2012; Marlina, 2009; Straker, 2016). This adds to my findings’ critique of the way inclusion is conceptualised, especially when it comes to participation and contribution. In the original Mor Barak Inclusion-Exclusion Scale, which was adapted for this research to the classroom setting, a question asked to participants included “I am typically involved and invited to actively participate...”. This leaves room for improvement in terms of clarifying what ‘involved in’ and ‘participation’ both mean. Participation must be expanded beyond just speaking. Students who do not speak and have high MBIE scores are prime examples of how low participation and contribution can still lead to a high perception of inclusion in the group.

Moreover, my findings uncover an argumentative discourse within the tutorial groups given the distribution of interruptions in this analysis. 66.67% of the top five speakers in the tutorials interrupted, while for all other speakers, just 18.18% interrupted others. This could potentially reflect language barriers too, in that non-native speakers may be less likely to interrupt as it takes a little bit longer to digest and respond to what is being said. Additionally, white speakers made up the majority of interrupters, and generally interrupted other white participants or the whole class. Knowing that interruptions occurred so frequently for white speakers is helpful information for tutors to monitor class discussions and attempt to mitigate the argumentative nature of the class discourse. In this way, it is worthwhile again to consider the literature surrounding international students in the UK and other European HEIs to understand that
international students are not simply lazy or passive, but rather choosing to not verbally engage for a number of reasons (Lomer & Mittelmeier, 2021). In acknowledging this, the university can help deconstruct the inequality regime in place, whether it is founded in race, nationality, or international student status.

To begin rectifying the inequality regime in place, the university should acknowledge it and prepare staff. This should include both high ranking administrators, as well as lower ranking educators such as tutors. As Tate & Page (2018) write, “[...] challenging the Racial Contract itself, it would require an acknowledgement of participation within systems of racism that privilege whiteliness” (150). The status of internationalisation, and the patterns of racial exclusion that potentially come with it, should be common knowledge for anyone who works at the university. For example, it should be well known that since new policies in 2013, the UK has pushed for internationalisation (Lomer & Mittelmeier, 2021). This directly affects who ends up sitting in the classroom as noted by the +10.23% increase in non-EU international students from academic year 2011/2012 to academic year 2021/2023. Notably too, even with internationalisation, the diversity in the classroom will be limited to those students who can sit exams with top scores and who can pay the noticeably higher international tuition (Bamburger et al., 2019). Even so, the university can better prepare educators for the students who will be present. The university can also take a stance by directly calling out to its educators that the pattern of exclusion in the classroom is not a reflection of ability, but rather a culmination of many reasons (Lomer & Mittelmeier, 2021). As from my own experience detailed above in facilitating the inclusion training, there are tutors who are not oblivious to these patterns, even while the university lacks communication to or preparation of educators. The university could even create a package of resources for tutors on helpful information for including non-EU international students more easily. There is power in being able to point students to helpful resources at the university, which may in turn help students overcome language barriers. All of this could be done to overcome the essentialist argument that the growing number of international students, and largely Asian international students, cannot and will not participate in the classroom (Marlina, 2009).

Another potential implication of the findings detailed above is that students may be getting stuck in patterns of inclusion and exclusion. The centrality measures highlighted that tutors act as controllers of the classroom discussion, just as Mameli et al. (2015) found too. The three tutors maintained the highest centrality scores for all four centrality measures, with the exception of in-degree in Tutorial A as the node receiving the most communication was the full class followed by the tutor. This is aligned with the literature in that tutors dominate tutorial discourse (Hardman, 2016). Importantly, on the university website’s explanation to students of what tutorials are, it lists that tutorials are primarily a place for all students to improve their communication
skills. This does not seem to be the case according to my findings though. Beyond tutors, it is common for only a couple students to drive tutorial discussions making it rare for all students to engage equally in conversations. Tutorial B shows how even in a classroom where most students participate orally and sit in a productive U-shape, a few students may still get stuck as isolates. They may require a bit more direct engagement from the tutor to increase their overall feelings of inclusion in the course. As controllers, tutors need to be aware of these students who are being isolated and feel isolated, to try to rectify the situation over the course of the ten weeks.

This once again emphasises a need for more professional development of tutors to engage in inclusive pedagogies. This brings up the question of how universities can support tutors’ professional development, which in turn, feeds into the student experience for all students. This is of course keeping in mind limited preparatory time given to tutors. One thing tutors could do is utilise student names. When the centrality measures were analysed along with the documentation of tutors name-calling, it was clear that it is uncommon for tutors to use a students’ name, and if they do, it is usually for students who are central to discussions already. It is firmly uncommon for tutors to use students’ names out of attendance check-in. From my own time tutoring, I know it is a struggle to remember names and pronunciations of names, especially for those students from cultures I am unfamiliar with. However, it is a small gesture that could promote higher feelings of inclusion. Hardman (2016) found that tutors check-in more with international students. If tutors are checking in more, but not using those students’ names, then how personal or genuine can that actually feel to the student? Universities could provide students with desk name plates to carry from class to class, or even financially compensate tutors with more preparatory time prior to the term beginning to familiarise themselves with their students’ names. This small gesture could prove to make a difference in students’ experiences of inclusion or exclusion.

Relating back to the idea of expanding the definition of participation in the classroom, the university could support tutors in implementing physical artefacts. University-wide training for tutors could include explicitly defining what participation in the classroom can look like. Participation could be defined as not just talking, but also listening, preparing work ahead of time, reading, reflecting, and critically thinking even if not vocalised. From there, the university could provide a tangible list of actions tutors can partake in to promote all of these, not just talking. In doing so, it could also help the tutor break down their facade as the “embodiment of knowledge” which may hinder students’ feeling safe to counter course learnings (Marlina, 2009, p. 236). Physical artefacts would help expand the definition of participation (Heron, 2019). Physical artefacts, such as pre-tutorial worksheets to complete allow students ample time to prepare and think of any questions that could be asked in the tutorial session (Heron, 2019). They offer an academic “scaffolding”, useful especially for students who may not
feel as comfortable with or confident in their English (Heron, 2019, p. 268). Instructional
design promotes learning and friendship ties in the classroom across nationalities
(Rienties et al., 2013). Of course, this kind of preparation by tutors would require the
university’s financial support and a vast change in their precarious contracts.

7.4: Limitations

A few limitations of this piece of research pertain to data collection. To begin,
there was only one recording device in the tutorial sessions. While the iPad had a
generally wide frame of view, participants were sometimes not stagnant. Tutors,
especially in Tutorials A and B, would move around the room walking in and out of
frame. Students, while less active, would sometimes lean back in their chairs or
accidentally move out of frame. While the audio of this would still be captured for the
most part, it was less high-quality. Furthermore, due to a technological malfunction, the
iPad, with a full battery, turned off in Tutorial B’s end of term session recording with still
about 10-15 minutes left of class time. To add to the limitations with data collection,
just three tutorial groups meant demographics that may be relevant to classroom
inequality regimes did not have large enough numbers to study. For instance, gender,
sexuality, disability, etc. could not be analysed in terms of social network analysis
without the possibility of putting a participant’s anonymity in danger. Furthermore, in
terms of data collection, four students never filled out demographic information. All four
were from Tutorial A, the tutorial in which I sat in on during all three recording sessions.
While I feel confident in the racial categories I assigned due to my interactions with
them, knowledge of their names and languages spoken, they never themselves
confirmed to me their race. End of term MBIE survey response rates were also quite low
at 54.05% for students in the tutorials observed, with Tutorial A having the lowest
number of just 4 students of 14 completing the second survey. This feeds into the
limitation of my study only involving just three tutorial groups for data collection, which
in turn adds to the difficulty in reflecting on how the findings apply to the larger
university. This was likely as students were overwhelmed with exam time, being
bombarded with emails pertaining to university-wide strikes, and/or focusing less on
academics and more on the World Cup tournament (as one tutor hypothesised).

Other limitations affected data analysis and coding in particular. For example,
speaking times were estimated in terms of seconds. This meant that, as an example, if
a student said “Yeah”, it was coded as one second even though it may not take the
student one full second to say. Physical communication, such as nodding, was not
coded for. Unfortunately, with just one iPad, everyone’s faces were not always visible,
especially if they faced away from the camera. As a result, other forms of
communication such as laughing and smiling were not taken into account. It should
also be mentioned that university policy meant that masks were not mandatory during this study. Some students still chose to wear masks making it difficult at times to fully understand what these students were saying, or for how long they were speaking. The use of technology in the classroom was also not taken into account. It could be argued that there is a big difference between a student speaking face-to-face to someone else, rather than speaking while staring at their phone, computer, tablet, or a piece of paper. All of this was not captured, but could prove worthwhile to study in future research.

Furthermore, in terms of data analysis and coding, tutor actions and participant interruptions present some limitations too. For example, the coding of interruptions may be seen as subjective. While I did my best to create a clear definition of what an interruption was and was not, it of course remains up to my own discretion. I did re-watch the videos and re-code the interruptions twice to ensure none were missed and that none were assumed to be interruptions if they were not. Additionally, I did not code at all for tutor actions that could be considered inclusive or exclusive, beyond interruptions, use of students’ names, or touching. Examples of actions that may have proven fruitful to code for could include a list of actions that were taught to tutors in one of the inclusion training sessions. To illustrate, in watching the recordings, I picked up on the fact that one of the tutors used the phrase “you guys” frequently. This was despite the tutorial session being predominantly females. Some tutors also alternated between using inclusive language like “let’s do this” or “we will do this”, rather than “you all will do this”. Further coding of class dialogue as questions, statements, and closed or open may have been fruitful too. Hardman (2016) found that students posing questions in tutorials is rare, and also found levels of closed and open questions are influenced by student demographics (Hardman, 2016). All of this could be seen as influencing the classroom dynamics and ultimately who felt invited into the discussions, and who felt and was left out. To overcome this, I could have also done 1:1 interviews and focus groups with the students and tutors to further understand how certain actions made them feel in terms of inclusion/exclusion. Unfortunately, as I was conducting data collection for both of my studies during the same term, I simply had too little time to do so.

7.5: Chapter Conclusion

This chapter focuses on the use of social network analysis to measure more covert behavioural manifestations of inclusion/exclusion within three tutorial groups. It may be seen as additional layer of analysis to Chapter 6, uncovering key insights that would be missed if the research only utilised survey methodologies. Over the course of one term, each tutorial is recorded three times. The dialogic data are visualised and analysed with common centrality measures. All together, the findings show that an
inequality regime within the classrooms studied exists, potentially according to race or international student status. Students who identify as BAME contribute and participate significantly less in the classroom, causing some of them to sit in the periphery of networks. Fortunately, this does not directly equate to low levels of perceived inclusion for these students. Even so, these findings present significant implications for inclusion researchers, as well as UK higher education institutions. This chapter ends with a review of the findings and a call to inclusion researchers to expand the definition of participation, which in turn affects how inclusion is conceptualised. It also discusses how UK higher education institutions must better educate and prepare tutors for facilitating tutorials with increasing numbers of international students. Social network analysis presents a clear opportunity for institutions to monitor covert levels of inclusion at the group-level utilising dialogic data. Without social network analysis, the main finding that there are significant differences in inclusion for some racial and ethnic minorities would be missed. This methodology may then be used to feedback into professional development for academics at all levels.

Chapter 8: Conclusion

In writing this dissertation, I had the goal of exploring social phenomena that are difficult to measure; namely, inclusion/exclusion and gender-based violence within one academic institution. I began my inquiry with a foundation in Acker’s (2006) inequality regimes. According to Acker (2006), each and every organisation will maintain a unique and usually covert social hierarchy. My underlying motivation throughout this process has been to uncover inequality regimes’ hierarchies after the implementation of organisational interventions, such as inclusion and gender-based violence trainings. Organisational inequality regimes reconstruct hierarchies pre-existing regionally and nationally along different salient dimensions of diversity (Tomaskovic-Devey & Avent Holt, 2019). This has put the larger context of academia within the United Kingdom at the forefront of my work, and subsequently at the core of my concluding musings too. I once again urge readers to take note of my delicate position as both an insider-outsider in committing this final chapter to paper.

My work maintains a theoretical contribution resulting from my reviewing of the literature in Chapter 2. I found that both the gender-based violence and the inclusion/exclusion literature are limited in their preoccupation with individual-level and overt displays of violence. This results in a fixation on people’s perceptions of violence, operationalised in research through survey mechanisms and high-level reporting measures. Certain researchers have urged for change and they inspired parts of what my new theoretical framework advocates for. Notably, I draw on Van Dijk & Khattab’s (2021) distinction between inclusion as a behaviour and a perception. I further build on
Banyard’s (2011) ecological model of bystander intervention by agreeing that the larger contexts of the exosystem and microsystem will help researchers move beyond their concern with the individual. In my theoretical contribution, I become a proponent of expanding research more into the group-level and covert manifestations of inequality regimes, and their resulting social phenomena such as gender-based violence and inclusion/exclusion. This change helps balance an overreliance on perceptions with a long-overdue examination of behaviours too. In putting forth this new theoretical framework, I find that there is a possibility for researchers, like myself, to borrow methodologies from the field of data science to propel the literature forward. With that, there is value in drawing methods and theoretical perspectives from other fields to expand the research that is being done on inclusion/exclusion and gender-based violence.

To achieve the above, I conducted two separate studies at one UK higher education institution. My first study utilised semi-structured focus groups and individual interviews, as well as surveys to understand how trust varies along with the implementation of a gender-based violence training intervention. To analyse these data, I employ thematic analysis, statistical analysis, as well as computational text analysis. I explore how the latter can be used to explore group-level covert perceptions of trust within participants’ dialogue, as a way to combat an overreliance on one’s own reporting of their experience while boosting reproducibility. My second study of inclusion/exclusion used surveys, observations, and recordings to conduct survey analysis and social network analysis. Once again, I demonstrated how computational social science methodologies may be used to augment the sociological study of these social phenomena if we permit ourselves to refocus on group-level and more covert manifestations. My methodological contribution is therefore showcased through my use of social network analysis and computational text analysis, with my work incorporating these more reproducible methods into inclusion/exclusion and gender-based violence research, respectively.

In this final chapter, I review my contributions, including my empirical findings too, while expanding on my final concluding thoughts across both studies. In particular, I offer a summary of my research’s implications and limitations. I also discuss how future research may build on the work I have done. I end this chapter, and this thesis, with a deliberation of how I believe the state of higher education has played a role in inclusion/exclusion and gender-based violence beyond this one field site. I’d like to acknowledge that my cynicism is palpable in this closing chapter. At the same time, I hope I have provided ample recommendations throughout this thesis, and below, that equally demonstrates my hope that transformation of academia is also possible.

8.1: Summary of Findings and Contributions
This section outlines the high-level findings of my research per each research question. I relate these findings back to the literature to highlight my contributions.

8.1.1: What variations in inclusion/exclusion are associated with the implementation of a diversity and inclusion training intervention within a university classroom?

Chapter 6 and Chapter 7 revealed several variations associated with inclusion after the implementation of a diversity and inclusion training intervention. While Chapter 6 was the first piece of research, to my knowledge, to adapt the Mor Barak Inclusion-Exclusion scale to the classroom context, Chapter 7 was the first, again to my knowledge, to incorporate social network analysis with demographic analysis to the classroom context. Through statistical analysis of the survey data, I failed to find evidence of any inequality regime present despite race and ethnicity being the most salient dimension of diversity for differences in perceived inclusion/exclusion. Chapter 7 echoed this finding to a certain extent as a covert inequality regime was in place whereby race, and potentially students’ subsequent national origin, were critical in regard to inclusion/exclusion in terms of actual participation and contribution in the classroom. Consequently, Chapter 7’s findings are at odds with Chapter 6’s findings and highlight how social network analysis offers more nuance into group-level more covert behaviours of inclusion/exclusion. Chapter 7 also indicates that white students dominate class discussions and administer the most amount of interruptions. Furthermore, students who at the periphery of class discussions tend to be Asian-identifying, while students at the core tend to be white-identifying.

Even as Chapter 7 found an inequality regime in place according to racial differences, Chapter 6 revealed that inclusion scores were significantly higher in Survey #2 later in the term than in Survey #1 earlier in the term for all racial and ethnic demographic groups. These findings go against Lomer & Mittelmeier (2021) who studied how educational research often discusses how international students, who generally are Asian-identifying, perceive western classrooms as hostile. My work is consistent with Van Dijk & Khattab’s (2021) theoretical work though in that exclusionary behaviours may be objectively happening, but they may not be interpreted as exclusionary or negative by the students. To add to this, my research highlights how lower levels of perceived inclusion tend to have to do with interactions with the tutors themselves, not with other students. This finding builds on Finkelstein et al. (2021) who writes how inclusive teaching must move to include inclusive behaviours beyond that of collaboration and communication. Even while a direct link between the findings and the training intervention cannot be made, this study shows that inclusion trainings for teaching staff cannot be the only inclusion intervention implemented. Both perceived
and observed behavioural exclusion were still present, indicating that further studies and interventions are needed.

These chapters indicate that we, as inclusive education researchers, are only at the beginning of understanding what inclusion is, and how it manifests within the academic context for students, especially taking into consideration the changing landscape of the ivory tower. My research has only investigated inclusion as participation and contribution, rather than also scrutinising uniqueness and belongingness too. Given this limitation, and the fact that an inequality regime, however covert, exists along racial and ethnic lines in the classrooms studied, there is still much more work to be done. I mentioned too in earlier chapters that some tutors have already demonstrated their knowledge of this inequality regime. This piece of research should only escalate their concern and render it a high-priority for the university itself. As described in Chapter 7 in particular, if an increase in international students is a proxy for an increase in racial minorities at the university, internationalisation needs to become a larger topic of conversation within inclusion discourse at the university.

A need to discuss the shifting environment of the university is even more advisable given recent university statistics. According to 2021/2022 university reports, 29.16% of the university’s undergraduate entrants were Black, Asian, and minority ethnic. This was just 7.8% in the academic year 2013/2014. Of these BAME students, 45.85% are Asian, with 64.3% of international students identifying as Asian. As I elaborate in Chapter 7, it appears the university, like many other UK universities, is continuing to pursue neoliberal logic by increasing the number of international students for financial gain (Bamberger et al., 2021). As Peter Mathieson, vice-chancellor of the University of Edinburgh succinctly stated, “We do rely on China- but so does every university” (Linklater, 2022). This puts the degree programmes solely relying on international tuition into a state of potential instability. Furthermore, if this is the case at all UK universities, universities must acknowledge potential patterns of racial and ethnic exclusion in the classroom; the patterns that some especially caring tutors on the front lines are already contemplating. Education of teaching staff at all levels could include directly discussing why these patterns of exclusion may exist beyond the false narrative of poor English and lower ability (Lomer & Mittelmeier, 2021). If the university is concerned with the wellbeing of all students, education of staff on what participation looks like is a crucial next step from this research. All together, this can fight the essentialist argument that international students of mainly Asian minority backgrounds do not want to participate in hostile UK classrooms (Marlina, 2009).

8.1.2: What variations in trust within a university’s student population are associated with the implementation of a gender-based violence training intervention?
Chapter 4 and Chapter 5 explore variations in trust as they relate to the implementation of the gender-based violence training intervention. While Chapter 4’s thematic analysis uncovers self-trust and social trust may have been boosted marginally by the training itself, Chapter 5 reveals how limited social trust may be. Furthermore, Chapter 5, through computational text analysis echoes Chapter 4’s finding that institutional trust is severely limited too. Potential reasons for the limited social trust, even despite the boost due to increased knowledge, awareness, and skills, may be due to social relationships. Students felt more at ease intervening for friends rather than strangers despite any heightened feelings of social repercussions they may experience, which is aligned with the previous work of Seo et al. (2022). At the same time, Chapter 4 revealed two critical barriers to trust in others to intervene, including gender dynamics and environmental factors. As Ermer et al. (2021) found, there is a double standard at play when it comes to trust in potential bystanders intervening. While students felt a heightened sense of self trust to intervene for same sex victims, social trust was limited by students’ internalised stereotypes of different levels of violence for the binary genders. Environmental factors aligned with Blayney et al. (2021) whereby self-trust and social trust were limited when scenarios involved it being dark out, students being alone, alcohol consumption, and students being in bars.

My findings in these two chapters provide evidence that gender-based violence trainings, just as inclusion trainings, are a necessary first step in counteracting gender-based violence. Yet, they cannot be a one-and-done fix. Firstly, this is because students post-training had a significantly higher desire than students pre-training for safety with their peers. Secondly, inequality regimes persist, in this case, for students from sexual and gender minorities. The trainings in a way may be contributing to universities avoiding further accountability. Given the university has yet to make this training mandatory, their gender-based violence strategy to date remains mediocre. By simply offering the training and not mandating it, the university may be seen to be “airbrushing” (Phipps, 2018, p. 230). At the time of writing this, it has been nine months since I delivered my final report to the university on the persistent inequality regimes post-training. Yet, the training remains voluntary and no further programming has been implemented to tackle the hurdles to social and institutional trust that I outlined in my research. This means the university has been made aware of how sexual and gender minorities continue to be at higher risk to gender-based violence than their majority group counterparts. Furthermore, my struggle with bureaucratic hurdles in attempting to secure incentives for students who do take the training shows an institutional preoccupation with financial gain, rather than the wellbeing of their end customer: the students. As Phipps (2018) writes, “The combination of gendered and intersecting structures within institutional airbrushing places the experiences and needs of survivors as secondary”. For this reason, I spend significant time in Chapter 4 and Chapter 5.
outlining steps the university could take to reconnect with the students as experts in their own experiences and to improve the wellbeing of all students, not just certain demographics of them. This included extra-university partnerships, mandating, and modifying the training with more inclusive sexual narratives, as well as implementing small-group cross-identity discussions to further establish an unspoken norm of bystander intervention.

8.1.3: Based on the proposed theoretical framework for gender-based violence and inclusion/exclusion, can a research blueprint be established for more holistically measuring trust and inclusion utilising computational social science methods?

In this thesis, I do establish a research blueprint for more holistically measuring gender-based violence and inclusion/exclusion. I begin this process in Chapter 2 by outlining how the literature devotes too much focus on individual-level and some group-level overt manifestations of these social phenomena. For instance, for gender-based violence in the higher education context, this includes reports of on-campus rapes and stalking. For exclusion, this may include admission statistics for certain demographic groups. Instead, I propose a new theoretical framework for understanding inclusion/exclusion and gender-based violence. This framework details how these social phenomena may manifest in overt and covert manners, and how they can be investigated at both the group-level and individual-level. My framework provides an opportunity for researchers to explore more group-level and covert mechanisms, especially mechanisms of social closure. In turn, this allows us to expand our toolsets beyond that of solely surveys to more reproducible computational social science. I specifically showcase how the group-level and covert manifestations of inclusion/exclusion and gender-based violence may be researched in Chapter 5 and Chapter 7.

In Chapter 5, I create a methodology for incorporating computational text analysis into gender-based violence research. Largely inspired by Alsaid et al.’s (2022) process, I provide a step-by-step overview for creating a bespoke domain-specific dictionary to measure social trust and institutional trust as it relates to gender-based violence. Notably, I show how even with a small training text corpus, bespoke dictionaries can easily be developed. To validate my findings, I ground my work in Chapter 4’s thematic analysis whereby I find institutional trust to be lower than social trust. Computational text analysis and specifically, sentiment analysis with a bespoke domain-specific dictionary, offers an additional and reproducible method of analysis for qualitative data. Accordingly, my method and the dictionary developed can be applied to larger data sets in the future where thematic analysis may prove too time-consuming.
Chapter 7 integrates social network analysis to add more nuance to my research by investigating group-level and more covert manifestations of inclusion/exclusion. This moves my work beyond that of perceptions to explore behaviours of inclusion/exclusion too. As mentioned, this chapter is the first piece of research, to my knowledge, that combines social network analysis with demographic analysis to the classroom context, all the while taking into consideration both the students and the educator. Therefore, my work is the first to combine both perceptions of inclusion/exclusion with behaviours of inclusion/exclusion for all those involved in the classroom milieu. The only previous mixed methods research, again to my knowledge, was Dawson (2008). Dawson (2008) studied student perceptions of inclusion in an online forum, but did not take into consideration the students’ identities and the potential salience of their identities in social relationships. In this chapter, I describe how social network analysis adds nuance to the survey findings, just as computational text analysis can be utilised in conjunction with thematic analysis and survey analysis. I accentuate how some key social network analysis metrics may prove useful if inclusion/exclusion continues to be conceptualised as participation and contribution.

8.2: Implications & Recommendations

My work here offers suggestive evidence that inequality regimes persist in terms of gender-based violence and inclusion/exclusion within my field site. On one hand, the depth of some of my findings are surprising to me when the university prides itself on values of being people-first and social responsibility forward. On the other hand, I am less surprised. Afterall, we cannot expect cookiecutter training interventions, such as the sexual consent and active bystander training discussed in this thesis, to relieve the specific inequality regime in place (Mellins et al., 2017; Githens, 2011). Furthermore, even when more individualised training interventions are implemented, we put a lot of faith in individual behavioural changes. As O’Connor et al. (2021, p. 8) writes, “It also usually assumes that once an alleged perpetrator is informed of their negative behaviour and the damage it is causing, they will undertake behavioural change: a problematic assumption”. Biesta (2023) expands on this whereby he likens educational interventions to pieces of art. In doing so, he highlights as educational researchers we pour every ounce of knowledge and passion we have into creating best practice interventions- our art. Real artists know better than to assume that their art will affect each person who views it in the exact same way or that the artwork will undoubtedly achieve their intended outcome. In short, we assume too much as researchers. Furthermore, as much as I feel I have personally advocated for these training interventions to not be a check the box activity, I fear they are. Anecdotally, nine months after first implementing the inclusion training, I met with an academic who also newly
maintained some EDI responsibilities. After running them through what the tutor inclusion training encompassed, they said to me they were excited to check this off their to-do list and inform their higher ups that they got it done. Apparently, months before, the academic had put forth ideas to senior members of the school for their EDI strategy campaign, including tutor inclusion training. However, at the time of them drafting their campaign, I had already implemented the tutor inclusion training. They did not ask about the training’s results at all. To me, this interaction poses the question: Are universities, and the decision-makers within them, just performing?

Performativity within academia is not new. Yarrow & Johnston (2021, p. 758) coined the term “institutional peacocking” to describe when higher education institutions employ micro-level changes, like training interventions, to boost reputational gain. By boasting these micro-level changes, HEIs are able to avoid any transformational change to the underlying status quo and oppressive structures in place (Yarrow & Johnston, 2021). Institutional peacocking allows institutions to continue to avoid accountability (Phipps, 2018; Yarrow & Johnston, 2021). This allows inequality regimes, exclusion, and gender-based violence to endure. Each of us, despite anyone’s best efforts, can get caught in the vicious cycle of performativity.

What is needed to move towards transformational change, and beyond performativity, is further action and reflection after these micro-level changes (Scott, 2020). I’ve touched on how the gender-based violence training intervention can expand its content to address some of the remaining inequality regimes and critical barriers that Chapter 4 discusses. I’ve also discussed changes tutors can make to their behaviours to foster inclusion. These are both within the realm of what I, myself, could likely implement at the micro-level of change. My biggest recommendation coming out of this piece of research beyond these micro-level changes is a desire for academia to pursue “deep, slow work” (Phipps, 2018, p. 239). This kind of transformational work involves listening to people’s experiences, taking accountability, and reconstructing the structures that bolster harm (Phipps, 2018). I join Scott (2020) and Forlin (2010) in their call for an institutional-wide approach to inequality regimes and their resulting violent social phenomena, like exclusion and gender-based violence. In doing so, I maintain a fear that transformational change will inadvertently create a byproduct of more bureaucracy, in true academic fashion. Researchers have shown evidence of this with the implementation of the Athena SWAN Equality Charter, whereby change comes in the form of large piles of paperwork usually at the hands of staff with marginalised identities (Yarrow & Johnston, 2021). In pursuing deeper work, academia must avoid a “neoliberal audit culture” and more document-obsessed procedures (Ahmed, 2012; Yarrow & Johnston, 2021, p. 765).

In pursuing deep work, academia must also take into consideration the larger historical and contemporary social context in which higher education institutions
operate. This thesis advocates for real change in the status quo. The first step of this must be people in positions of power to be reflexive; acknowledging what status quo and social hierarchies are in place. Reckoning with the privilege, entitlement, elitism, violence, and inherent exclusion that got higher education institutions this far is critical (Forlin, 2010). Furthermore, academia has co-opted a neoliberal agenda and higher education should not ignore that (Tight, 2019). I would like to avoid utilising the term neoliberalism as a catch-all though, and so I encourage decision-makers to figure out which aspect of neoliberalism is hurting those within academia. It could be globalisation, entrepreneurialism, colonialism, academic capitalism, marketisation, managerialism (and resulting paper pushing discussed above), massification, and/or privatisation (Tight, 2019). Without reflecting on these mechanisms at play in the larger social context, any policies or interventions will be rendered ineffective (Armstrong et al., 2011). One way to begin reckoning with inequality regimes is by pursuing a whole-school approach of resistance and coalition building (Shahjahan, 2014; Woods et al., 2022). By embracing the local communities, neighbourhoods, and the larger academic context, decision-makers in academia can become more conscious of how behaviours currently contribute to violence and inequality regimes (Shahjahan, 2014; Woods et al, 2022). I firmly believe these steps are the only way forward if higher education institutions are to genuinely mitigate gender-based violence and exclusion.

8.3: Limitations & Future Research

I’ve discussed limitations throughout my work, and instead here would like to express how limitations could guide future research. Specifically, I wish to discuss future research in a way that would adopt the above reflexive whole-school approach. First, one limitation to both my studies was that the cohorts I studied were primarily concerned with first year undergraduate students. My gender-based violence work focused on first year students living in student accommodation, while my inclusion study worked with mostly first year students in two schools only. As a result, my findings are limited to first year students, their experiences and behaviours of gender-based violence and inclusion/exclusion. In the future, I would instead advocate for embedding similar research into the larger school.

A few changes could be made to achieve this across both studies. For instance, in the inclusion training, it would be worthwhile to increase the sample size. This could include making participation in the study part of the entry-level course curriculums across schools. This does not have to include the recording and observations, which can still be optional. Rather, the study and the survey mechanisms could be explained to the students prior to signing up for the course. When they choose to enrol, they can confirm their consent or not. This way, students can be allocated to tutorial groups.
accordmg to their participation. Likely, if it were framed as embedded in the course from the get-go, this would also help boost participation numbers. In boosting participation numbers, other demographic groups less central to my work here can be explored. This could include more of a focus on students with disabilities, students with English as a second language, commuter students, etc. A larger sample size would also aid in fostering a more intersectional analysis. For the gender-based violence study, I recommend the university mandate the training for all first year undergraduate students.

As discussed, I firmly believe increasing incentives, which the university is unlikely to agree to anyways, would not automatically boost participation numbers or response rates. My experience with the two studies showed how, even with incentives, students do not wish to participate and become overwhelmed quickly. Part of me wonders if research on student populations is just difficult at this moment in time because of “post-COVID” fatigue. Especially for students in their first year of university, the autumn term represents learning to juggle much more than they likely have had to juggle alone in the past. To reach a wider sample of students, the trainings must be embedded in the first year onboarding process. In the same way as just discussed, a larger sample size could help highlight the nuances in different identity groups’ experiences and behaviours. Only with a larger sample size could the potential wider school inequality regimes be explored.

A second limitation is that my studies could not address how inequality regimes exist outside of the university community. Students do not live in a silo during their time at higher education institutions. They roam off-campus, they explore their cities, and they meet people not affiliated with the university; all of this still impacts their university experience. The inclusion study is even more limited in this way by solely looking at students’ experiences of one classroom. To counteract this, I would recommend future research incorporates coalition building to reach and better understand students’ interactions with the wider community. With that, research should recruit at least two groups of stakeholders from campus organisations and non-university related organisations to adapt both training interventions. Students would lead these adaptations as they are the end consumers of the training interventions. Expanding survey analysis to include more questions about students’ experiences beyond the boundaries of the classroom could help illuminate how inclusion/exclusion shifts during students’ university experience. I also recommend that future research intentionally recruits student leaders who identify as parts of the groups that are currently marginalised according to the inequality regimes discovered in this work. For example, the gender-based violence training adaptation coalition must include students from gender and sexual minorities. The inclusion adaptation likewise must include students from Asian backgrounds, especially those who are international. Without partnership
with the wider university community, intentional coalition-building, and an intersectional focus, future research will similarly lack an understanding of how inequality regimes at university persist.

8.4: Epilogue

I would be remiss to end this thesis without mentioning the current state of UK academia that is affecting students, staff, and those who teeter between the two, like me. When I began writing this document in February 2023, there was evidence that academia in the UK was on the verge of mayhem. I mentioned in an earlier chapter that my autumn 2022 data collection was affected by nation-wide strikes planned by the University College Union (UCU). By April 2023, we entered a full-blown marking and assessment boycott that continued until September 2023. 145 higher education institutions participated in this dispute against pay and working conditions. Further strike action is already planned.

These unfair pay and working conditions have personally affected me just like most other PhD students I know who live on precarious contracts. In May 2023, I was offered a short-term position to implement an inclusion survey for a new diversity, equity, and inclusion training at my home university. Despite it being weeks into the boycott, I was told I could not be paid monetarily, but that the university could pay me in high street vouchers. Sadly, vouchers to Boots do not pay rent nor will they pay for bus tickets to campus. There is a certain kind of irony to being asked to do diversity, equity, and inclusion work without equitable pay. This is all the while the vice-chancellor of my university, Peter Mathieson as previously mentioned, received a £43,000 bump to bring his annual salary up to £406,000 (Ross, 2023). According to the UCU (2021), other perks of the job include two 24-hour chauffeurs and rent for a five-bedroom house. The marking boycott and pay dispute meant about 2,000 students at my institution graduated in July without any final marks (McCool & Godden, 2023). Similar experiences are being had across the UK, as evidenced by the participation of 145 higher education institutions in the marking and assessment boycott. All this to say, I deeply want to conclude this thesis by stating that UK higher education is failing students despite the best efforts. Unfortunately, I am fairly certain these are not any administration’s best efforts. Best efforts are being absolutely stifled by a gross preoccupation with financial gain and performativity in UK academia. That has been made clear to me.

Despite my pessimism, I maintain hope that transformational change can happen. I likewise sincerely hope this thesis has helped show how group-level and covert manifestations of inclusion/exclusion and gender-based violence must continue to be studied. I’ve demonstrated here that with the right mixture of tools, we can

207
finetune training interventions to ensure they are effective first steps towards better futures in academia and beyond. Yet, as of right now, violence is in the air we breathe and it is suffocating some more than others. How lovely would it be though if those in positions of power relinquished their power, let go of neoliberal logic, and embraced coalitions with those with less power so we can fix it?
Appendix

Appendix A: Additional Field Site Context

A.1: Existing Strategies for Mitigating Gender-Based Violence & Fostering Inclusion

Before the start of the study, the university had an extensive equality, diversity, and inclusion strategy, but it only had a preliminary gender-based violence mitigation plan. As of 2022, students and staff alike could volunteer to take the Epigeum “Consent Matters” and “Tackling Harassment” training. The training was a result of a student-led bottom-up campaign for consent modules. Despite limited staff headcount, the university did recruit an entire team as of 2023 devoted to mitigating gender-based violence and an ample number of partnerships with gender-based violence organisations both on- and off-campus. Student and staff training for equality, diversity, and inclusion was much more decentralised and offered through individual university offices (ex. Office of Accessibility), but remained mainly voluntary. The only mandatory D&I training at a university-wide level was on unconscious bias and only mandated for staff involved in hiring and workforce planning. The same training was available for students on a voluntary basis. Students and staff also had access to LinkedIn Learning where several D&I-related trainings were suggested. Additionally, ad hoc events through staff diversity networks supplemented university-provided learning opportunities. Mandated D&I training for teaching staff was highly school dependent within the university. While the D&I trainings for tutors put together for this study were mandated, they were the first of their kind and no equivalent training was mandatory for any other level of teaching staff. This was also exclusive to the two schools from which tutors were recruited, plus one other school which had no undergraduate programme at the time.

Appendix B: Gender-Based Violence Study

B.1: Gender-Based Violence Study Research Design per Cohort

<table>
<thead>
<tr>
<th>Research Stage for Participants</th>
<th>Cohort 1</th>
<th>Cohort 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Express interest in the study via email or QR code signup</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Sign the consent form after eligibility for study confirmed</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Complete pre-training survey</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Complete Epigeum training</td>
<td>X (Consent Matters module only)</td>
<td>X (Consent Matters &amp; Tackling Harassment modules)</td>
</tr>
<tr>
<td>Complete focus group or online survey equivalent</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Complete post-training survey (one week or three weeks post-training)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Complete post-training individual 1:1 interviews</td>
<td>X</td>
<td>N/A</td>
</tr>
</tbody>
</table>
B.2: Gender-Based Violence Study Consent Forms

Consent & Active Bystander Training Evaluation Research Study
Participant Information Sheet (Cohort 1)

We would like to invite you to take part in a research study. Before you decide whether you want to take part, you should understand why the research is being done, and what it would involve for you. Please take time to read the following information carefully and contact us if you need further information before deciding whether or not you wish to participate.

What is the purpose of the study?
The aim of the project is to evaluate the university’s Epigeum active bystander and consent training, to establish and identify factors which contribute to bystander training being effective in delivering outcomes, and to consider what additional resources will be needed to make the current and future university’s training suite more effective in creating a safe environment that does not tolerate discrimination and harassment in any form, including gender-based violence. The student surveys and focus groups are critical to this process of evaluating the university’s training programme.

Why have I been asked to take part?
You have been asked to take part because you are an undergraduate student currently living in university accommodation.

If I agree to participate in the study, what will I take part in?
- **Training (~ 45 minutes total):** You will be asked to complete the university’s bystander and consent virtual training. This training focuses largely on identifying and mitigating gender-based violence and sexual harassment. This training, while self-paced, should take no longer than 45-minutes to complete.
- **Surveys (~ 1 hour total):** You will be asked to complete two surveys focusing on your knowledge, behaviours, feelings of inclusion, and awareness of gender-based violence and sexual harassment. One survey will be before the training and one survey will be after the training occurs. These surveys will take no longer than 30-minutes each to complete.
- **Focus Groups (~1.5 hour total):** You will be asked to participate in one 1.5 hour focus group with other members of your resident hall. This focus group will either be before or after you take the training, and be either in person or online via Teams (Covid-19 dependent). The primary PhD researcher, with the assistance of the research intern, will facilitate these sessions with half the time spent as a traditional focus group, and half the time spent on a hypothetical situation. For the focus group portion of the session, it will be based on questions seeking to evaluate changes in awareness, knowledge, and behaviours.
• **Written comments**: If you would like to give your views but do not want to participate in a group interview, we will send a link to an anonymised survey with a single free text space for you to complete.

**What are the possible benefits and disadvantages of taking part?**
We expect that results from the study will help us to evaluate the university’s current bystander and consent training, as well as guide future diversity, equality, and inclusion strategy. If you participate in the interview, you will contribute some of your personal time. Discussing issues relating to gender-based violence and sexual harassment can be triggering and highly sensitive; we will try to ensure participants do not feel uncomfortable during the group interview. We will also remind students throughout the study of university support and resources that you can access, including such tools as the [link redacted] platform. If you decide to be involved with this research, the team will provide further materials you can use for support as well.

**Do I have to take part?**
It is entirely up to you whether you decide to participate in the research or not. If you agree to participate, you are free to withdraw at any time and without giving a reason. If you participate in the focus group and surveys, but then decide you want to withdraw, your data (responses to surveys) will be deleted as long as you let us know before the data have been analysed/written up. Deciding not to take part, or changing your mind at a later stage will not affect the education or support that you receive now or in the future.

**What happens to my information?**
Your data will be processed in accordance with Data Protection law. All personal information you provide during the course of the research study is kept confidential. The focus group will be recorded, transcribed (typed-up) and analysed. Any direct identifiers or information that might identify you will be removed (ex. Residential hall name, club involvement, etc.). Instead, students will be given numbers. While names will be retained initially to help link the surveys, focus groups, and social network analysis portion of analysis, once the data collection is complete, names will be completely replaced by unidentifiable numbers. All data with names will be destroyed. Indirect identifiers such as demographic information will be retained, but will not be shared at an individual-level. Rather, demographics will be used to extract high-level programme-wide analysis based on gender, race, age, etc. We will follow ethical and legal guidelines and keep all information privately and securely. Any personal stories of GBV shared in the surveys, focus groups, or training with the research team will not be shared publicly unless they are stripped of personal identifiers. All documents will be stored electronically in a secure [name redacted] space. Anonymised data will be stored for a maximum of 5 years. While data is stored, it will be retained on a password protected server.

**What will happen to the results of the study?**
The results of the study will be used to develop a best practises and literature review document to be shared with the university. The results will be published in a doctoral
thesis, as well as in gender-based violence journals. You will not be identifiable in any published results, and we will make the papers available to students after publication.

How do I take part?
If you are interested in taking part, please sign the following consent form and send to [name redacted] at [email redacted].

Complaints
If you have any concerns or complaints about this research, please contact the [name redacted] at [email redacted].

Data Protection
For general information about how we use your data please go to: [link redacted]
Alternatively, please contact the [name redacted]: [link redacted]

Many thanks for your time and help

Consent & Active Bystander Training Evaluation Research Study
Consent Form

Please initial each box to confirm that you agree with the statements below. The statements will be re-read to you ahead of data collection and you will be asked to state your agreement.

<table>
<thead>
<tr>
<th>Please initial box</th>
</tr>
</thead>
<tbody>
<tr>
<td>I confirm that I have read and understand the participant information sheet for the above study.</td>
</tr>
<tr>
<td>I confirm that I have had the opportunity to consider the information, ask questions and have had these questions answered satisfactorily.</td>
</tr>
<tr>
<td>I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason and without any care or support and/or legal rights being affected.</td>
</tr>
<tr>
<td>I understand that data collected about me during the study will be converted to anonymised data and a number will be used to protect my identity. Any direct identifiers will not be shared, while my demographics will only be used for high-level programme-wide analysis.</td>
</tr>
</tbody>
</table>
We would like to invite you to take part in a research study. Before you decide whether you want to take part, you should understand why the research is being done, and what it would involve for you. Please take time to read the following information carefully and contact us if you need further information before deciding whether or not you wish to participate.

What is the purpose of the study?
The aim of the project is to evaluate the university’s Epigeum active bystander and consent training, to establish and identify factors which contribute to bystander training being effective in delivering outcomes, and to consider what additional resources will be needed to make the current and future university’s training suite more effective in creating a safe environment that does not tolerate discrimination and harassment in any form, including gender-based violence. The student surveys and focus groups are critical to this process of evaluating the university’s training programme.

Why have I been asked to take part?
You have been asked to take part because you are an undergraduate student currently living in university accommodation.

If I agree to participate in the study, what will I take part in?
- **Training (~ 5 hours total):** You will be asked to complete the university’s bystander and consent virtual training. This training focuses largely on identifying and mitigating gender-based violence and sexual harassment. This training, while self-paced, should take no longer than 5-hours to complete.
• **Surveys (~ 1.25 hours total):** You will be asked to complete two surveys focusing on your knowledge, behaviours, feelings of inclusion, and awareness of gender-based violence and sexual harassment. One survey will be before the training and one survey will be after the training occurs. These surveys will take no longer than 30-45 minutes each to complete.

• **Focus Groups (~1.5 hour total):** You will be asked to participate in one 1.5 hour focus group with other members of your resident hall. This focus group will either be before or after you take the training, and be either in person or online via Teams (Covid-19 dependent). The primary PhD researcher, with the assistance of the research intern, will facilitate these sessions with half the time spent as a traditional focus group, and half the time spent on a hypothetical situation. For the focus group portion of the session, it will be based on questions seeking to evaluate changes in awareness, knowledge, and behaviours.

• **Written comments:** If you would like to give your views but do not want to participate in a group interview, we will send a link to an anonymised survey with a single free text space for you to complete.

*What are the possible benefits and disadvantages of taking part?*

We expect that results from the study will help us to evaluate the university’s current bystander and consent training, as well as guide future diversity, equality, and inclusion strategy. If you participate in the interview, you will contribute some of your personal time. Discussing issues relating to gender-based violence and sexual harassment can be triggering and highly sensitive; we will try to ensure participants do not feel uncomfortable during the group interview. We will also remind students throughout the study of university support and resources that you can access, including such tools as the [link redacted] platform. If you decide to be involved with this research, the team will provide further materials you can use for support as well.

*Do I have to take part?*

It is entirely up to you whether you decide to participate in the research or not. If you agree to participate, you are free to withdraw at any time and without giving a reason. If you participate in the focus group and surveys, but then decide you want to withdraw, your data (responses to surveys) will be deleted as long as you let us know before the data have been analysed/written up. Deciding not to take part, or changing your mind at a later stage will not affect the education or support that you receive now or in the future.

*What happens to my information?*

Your data will be processed in accordance with Data Protection law. All personal information you provide during the course of the research study is kept confidential. The focus group will be recorded, transcribed (typed-up) and analysed. Any direct identifiers or information that might identify you will be removed (ex. Residential hall name, club involvement, etc.). Instead, students will be given numbers. While names will be retained initially to help link the surveys, focus groups, and social network analysis portion of analysis, once the data collection is complete, names will be completely
replaced by unidentifiable numbers. All data with names will be destroyed. Indirect identifiers such as demographic information will be retained, but will not be shared at an individual-level. Rather, demographics will be used to extract high-level programme-wide analysis based on gender, race, age, etc. We will follow ethical and legal guidelines and keep all information privately and securely. Any personal stories of GBV shared in the surveys, focus groups, or training with the research team will not be shared publicly unless they are stripped of personal identifiers. All documents will be stored electronically in a secure [name redacted] space. Anonymised data will be stored for a maximum of 5 years. While data is stored, it will be retained on a password protected server.

*What will happen to the results of the study?*
The results of the study will be used to develop a best practises and literature review document to be shared with the university. The results will be published in a doctoral thesis, as well as in gender-based violence journals. You will not be identifiable in any published results, and we will make the papers available to students after publication.

*How do I take part?*
If you are interested in taking part, please sign the following consent form and send to [name redacted] at [email redacted].

*Complaints*
If you have any concerns or complaints about this research, please contact the [name redacted], [name redacted] at [email redacted].

*Data Protection*
For general information about how we use your data please go to: [link redacted]
Alternatively, please contact the [name redacted]: [link redacted]

Many thanks for your time and help

Consent & Active Bystander Training Evaluation Research Study Consent Form

Please initial each box to confirm that you agree with the statements below. The statements will be re-read to you ahead of data collection and you will be asked to state your agreement.

<table>
<thead>
<tr>
<th></th>
<th>Please initial box</th>
</tr>
</thead>
<tbody>
<tr>
<td>I confirm that I have read and understand the participant information sheet for the above study.</td>
<td>☐</td>
</tr>
<tr>
<td>Please initial box</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>---</td>
</tr>
<tr>
<td>I confirm that I have had the opportunity to consider the information, ask questions and have had these questions answered satisfactorily.</td>
<td>□</td>
</tr>
<tr>
<td>I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason and without any care or support and/or legal rights being affected.</td>
<td>□</td>
</tr>
<tr>
<td>I understand that data collected about me during the study will be converted to anonymised data and a number will be used to protect my identity. Any direct identifiers will not be shared, while my demographics will only be used for high-level programme-wide analysis.</td>
<td>□</td>
</tr>
<tr>
<td>I agree to my interview being audio recorded.</td>
<td>□</td>
</tr>
<tr>
<td>I agree to my anonymised data being used in future studies.</td>
<td>□</td>
</tr>
<tr>
<td>I understand that all data will be stored in accordance with the [policy name redacted].</td>
<td>□</td>
</tr>
<tr>
<td>I agree to participate in the above study.</td>
<td>□</td>
</tr>
</tbody>
</table>

B.3: Gender-Based Violence Study Recruitment Communications & Setbacks

The study was pushed to students via social media with the university’s resident halls team, through resident advisors, in student accommodations with posters, and via a couple student unions. For Cohort 2 student recruitment, the main student union at the university also helped disseminate recruitment materials. Once a student expressed interest, I ensured they were eligible for participation. This meant I checked that they were actively living in student accommodation and identified as a first year student (part-time or full-time). When communications did not include the consent form itself, a QR code was used to redirect the student to a sign-up form. Once they signed up, they were sent the consent form directly via email. At least two follow-up emails were sent to
the potential participants who expressed interest, either by myself or the research assistant at the time.

The study’s initial design encountered several setbacks and subsequent alterations. The ethics approval process took longer than anticipated, with submission in December 2021, but approval only in February 2022. For this reason, recruiting for this study only began in late March 2022. The slow participant uptake extended the recruitment process, meaning the study began data collection in early May 2022. This study also took place while COVID-19 continued to affect the day-to-day for higher education institutions and students alike. It is possible that students were hesitant to sign up for in-person studies given the spikes in COVID-19 cases in early 2022. Once enough participants signed up, the earliest focus groups and training times could occur, unfortunately, coincided with students’ examination schedules, end of year socials, moving out of student accommodations, and summer travelling. Given the amount of juggling between academic, professional, and social activities students would need to do to fully participate in this study, the drop off rates throughout the initial study were higher than anticipated.

Early in the initial recruitment process, it was decided that this study should be replicated with a second cohort in the autumn 2022 term. Even though the second cohort would take place during a less academically intense time for students, a higher incentive for Cohort 2 students was still approved to attempt to recruit the greatest number of students possible. For the spring 2022 Cohort 1, students were offered catered food for the in-person focus group sessions. If they fully participated in the study, they were entered into a raffle for £50 for a student organisation of their choice. Cohort 2’s incentives increased to include catered food for in-person focus group sessions, and a raffle with the following prizes: an iPad grand prize, and three £100 vouchers to Asda as secondary prizes. It was the researchers’ hope that the higher incentives would be more worthwhile for students to participate in the full study.

The incentives also needed to be re-evaluated for Cohort 2 given that the training ended up taking longer than anticipated. Cohort 1 only completed the Consent Matters training module, and were not asked to complete the significantly longer Tackling Harassment training module. This is because the Consent Matters training was meant to take ~ 45 minutes, but it took some Cohort 1 participants up to 1hr 30 minutes. The Tackling Harassment training is meant to take two hours according to Epigeum. This means that the total training time for Cohort 2 was closer to 3.5-5 hours total depending on how quickly students finished the two courses. The researchers did not feel they could ask for participants to complete both courses for the Cohort 1 incentives given the already higher than anticipated participation drop off and lack of time allocated to the focus group sessions. The final research design was approved in April 2022.
B.4: Gender-Based Violence Study Survey Items

Pre-Training Survey

Thank you for taking this survey and participating in our research. This survey consists of 100 questions and should take about ~45 minutes to complete in one-sitting. As a reminder, these student surveys are critical to the process of evaluating the university’s bystander & consent training programme, while also highlighting where the university’s training suite could be more effective in creating a safe environment that does not tolerate discrimination and harassment in any form.

Trigger warning: Discussing issues relating to gender-based violence and sexual harassment can be triggering and highly sensitive, thus if you need to take a break between questions to collect yourself, we encourage you to do so. We also encourage you to utilise university support and resources, which you will find links to at the end of this survey, as well as in a PDF in the email this survey was sent in.

Your data will be processed in accordance with Data Protection law. All personal information you provide during the course of the research study is kept confidential. Indirect identifiers such as demographic information will be retained, but will not be shared at an individual-level. Rather, demographics will be used to extract high-level programme-wide analysis based on gender, race, age, etc. We will follow ethical and legal guidelines and keep all information privately and securely.

1) What is your participant code number (#)? If you are unsure, please check your initial participant information email or ask the researcher if you are taking this survey in person: __________
2) What is your age?: ________________________________
3) What is your sex?: Male, female, non-binary, trans male, trans female, I do not identify as any of these and instead identify as: ___________
4) What is your sexual orientation?: Heterosexual, homosexual, bisexual, pansexual, asexual, lesbian, gay, queer, I do not identify as any of these and instead identify as ___________
5) Which school are you a part of?: [Drop-down options]
6) Which programme are you in?: _______________________________________
7) Are you a member of a university-sponsored activity? (ex. Societies, academic competitions, sporting teams, university bands, student government, etc.): Yes, no
8a) What is your relationship status?: Single, monogamous, casually dating one person, casually dating multiple people, polyamorous, o I do not identify as any of these relationship statuses and instead define my relationship status as: ________________
8c) How many months have you been in a relationship? (If you are in multiple relationships currently, please put the number of months of your longest current relationship): ________________________

Post-training survey included all of the same questions, minus the demographic questions.
9) How many units of alcohol do you drink each week, on average? (e.g. 1 pint of beer is approximately 2 units, and one small glass of wine/one shot of liquor is 1 unit): 0 units, 1-7 units, 8-14 units, 15-21 units, 22-28 units, 29-35 units, 36-69 units, 70+ units

10) On average, how many days per week do you go out for a 'night out'? (e.g. party, social, bar, club, etc.): 0 days, 1 day, 2 days, 3 days, 4 days, 5 days, 6 days, 7 days

11) What is your race and ethnicity? Please note, you are able to select multiple answers here or write-in your own race & ethnicity if you do not see your’s below.: White, White-Scottish, White-Other British, Black or Black British-Caribbean, Black or Black British-African, Other Black Background: ____, Asian or Asian British-Indian, Asian or Asian British-Pakistani, Asian or Asian British-Bangladeshi, Chinese, Other Asian Background: ______, Mixed-White and Black Caribbean, Mixed-White and Black African, Mixed-White and Asian, Latina/Latino/Latino, Arab, Gypsy or Traveller, Not known, Information Refused, Other racial & ethnic background: _______

12) Have you previously taken any courses or trainings on being an active bystander?: Yes, no, I am unsure or do not know

13) Have you previously taken any courses or trainings on sexual consent?: Yes, no, I am unsure or do not know

14) Would you consider yourself a victim and/or survivor of unwanted sexual contact with someone?: Yes, no, I am unsure or do not know

15) Please read the following list of statements and select to what extent you feel you disagree or agree. [1- Completely disagree, 2- Somewhat disagree, 3- Neither disagree nor agree, 4- Somewhat agree, 5- Completely agree]

- 15) I feel I have a strong awareness and knowledge surrounding what constitutes sexual harassment.
- 16) I feel I have a strong awareness and knowledge surrounding what constitutes gender-based violence (GBV).
- 17) I understand what the potential consequences are for students who commit sexual harassment and gender-based violence (GBV).
- 18) I feel that the [name redacted] should make consent and active bystander training mandatory for all students.
- 19) I understand what university resources and support are available if I or someone I know is a victim and/or survivor of sexual harassment and gender-based violence (GBV).
- 20) If myself or a friend wanted to report an incident of gender-based violence (GBV) and/or sexual harassment, I would know how to report it to the relevant university services.
- 21) If I find myself a victim of unwanted sexual activity with a fellow student bystander present, I feel like I can expect the bystander to help me.
- 22) If I find myself a victim of unwanted sexual activity with a fellow student bystander present, I feel like there is a culture of unspoken permission for the bystander to help me.
- 23) I feel a sense of community with other people in my student accommodation (for example, you share interests and concerns with them).
24) It is very important to me to feel a sense of community with people in my student accommodation.
25) Some people care a lot about the kind of student accommodation they live in. For others, student accommodation is not important. What the student accommodation is like is very important to me.
26) I feel a sense of community with other people at my university (for example, you share interests and concerns with them).
27) It is very important to me to feel a sense of community with people at my university.
28) Some people care a lot about the kind of university they go to. For others, the university is not important. What the university is like is very important to me.
29) I feel a sense of trust with other people in my student accommodation.
30) It is important to me to feel a sense of trust with the people in my student accommodation.
31) I feel a sense of safety with other people in my student accommodation.
32) It is important to me to feel a sense of safety with the people in my student accommodation.
33) I feel a sense of trust with other people at my university.
34) It is important to me to feel a sense of trust with the people at my university.
35) I feel a sense of safety with other people at my university.
36) It is important to me to feel a sense of safety with the people at my university.
37) For each of the following statements, rate your degree of confidence in completing each task on a scale of 0 (not at all confident) to 100 (completely confident). Note, you must individually click into each slider and move it (even if you move it back to 50) for your response to be answered. [0- Not at all confident in doing, 25- Somewhat]
unconfident in doing, 50- Neither unconfident nor confident in doing, 75- Somewhat confident in doing, 100- Completely confident in doing]

- 37) Express my discomfort if someone makes a sexualised joke about a person's body.
- 38) Express my discomfort if someone says that rape victims are to blame for being raped.
- 39) Call for help (i.e. call 999) if I hear someone in my student accommodation yelling “help.”
- 40) Talk to a friend who I suspect is in an abusive relationship.
- 41) Get help and resources for a friend who tells me they have been a victim of unwanted sexual activity.
- 42) Ask a stranger who looks very upset at a party if they are ok or need help.
- 43) Offer to a friend to walk them home from a night out.
- 44) Ask a stranger if they need to be walked home from a night out.
- 45) Speak up in class if a professor is providing misinformation about sexual assault.
- 46) Criticise a friend who tells me that they engaged in sexual activity with someone who was passed out or who did not give consent.
- 47) Do something to help a very drunk person who is being taken elsewhere, away from their friends, by a person or group on a night out.
- 48) Do something if I see a marginalised (ex. Ethnic minority, woman, queer, etc.) person surrounded by a group of men at a party who looks very uncomfortable.
- 49) Seek help for someone else if I hear of an abusive relationship in my student accommodation.
- 50) Tell a Resident Assistant or other campus authority about information I have that might help in an ongoing sexual assault case even if pressured by my peers to stay silent.
- 51) Speak up to someone I know who is making excuses for forcing someone to have sexual relations with them.
- 52) Speak up to someone I know who is making excuses for having sexual relations with someone who is unable to give full consent.
- 53) Speak up to someone I know who is making excuses for using physical force in a relationship.
- 54) Speak up to someone I know who is calling their partner names or swearing at them.
- 55) Speak up to someone I do not know who is making excuses for forcing someone to have sexual relations with them.
- 56) Speak up to someone I do not know who is making excuses for having sexual relations with someone who is unable to give full consent.
- 57) Speak up to someone I do not know who is making excuses for using physical force in a relationship.
- 58) Speak up to someone I do not know who is calling their partner names or swearing at them.
59) Please read the following list of statements and select to what extent you feel you are likely to engage in the behaviours using the following scale: [1- Not at all likely to do, 2- Somewhat unlikely to do, 3- Neither unlikely nor likely to do, 4- Somewhat likely to do, 5- Completely likely to do]

- 59) Confront someone I know who plans to give someone alcohol to get sex.
- 60) Confront a friend if I hear rumours that they forced sexual activity on someone.
- 61) Check in with my friend who looks drunk when they go to a room with someone else at a party.
- 62) Say something to my friend who is taking a drunk person back to their room on a night out.
- 63) Confront a friend who is hooking up with someone who is very drunk.
- 64) Challenge a friend who made a sexist joke.
- 65) Report a friend that forced someone to engage in sexual activity.
- 66) Ask for verbal consent when I am intimate with my partner, even if we are in a long-term relationship.
- 67) Stop sexual activity when asked to, despite still being sexually aroused.
- 68) Stop engaging in sexual activity with a partner if they say to stop, even if it started consensually.
- 69) Decide not to engage in sexual activity with a partner if they are intoxicated.
- 70) Refuse to participate in activities where others’ appearances are ranked/rated.
- 71) Express my concern if a family member makes a sexist joke.
- 72) Call for help (ie. call 999) if I saw a group of men bothering a marginalised (ex. Ethnic minority, woman, queer, etc.) person in a public space.
- 73) Call for help if I saw a person that I do not know go to their student accommodation with a group of guys and hear the person yelling for help.
- 74) Go with a female friend to the police department if she says she was sexually assaulted.
- 75) Go with a male friend to the police department if he says he was sexually assaulted.
- 76) Go with a non-binary friend to the police department if they say they were sexually assaulted.
- 77) Visit a website to learn more about sexual violence.
- 78) Join an organisation that works to stop rape and sexual abuse.
- 79) Participate in a rally on campus to stop rape and sexual abuse.
- 80) Take a class to learn more about sexual violence and abuse.

81) Please provide a numerical answer for the following questions, doing your best to estimate (ex. 0 times, 1 time, 6 times, 10 times, etc.)

- 81) Approximately how many times this academic year have you spoken with friends about gender-based violence / sexual harassment? ________
• 82) Approximately how many times this academic year have you spoken with fellow student accommodation residents about gender-based violence / sexual harassment? __________

• 83) How many times have you spoken to friends about gender-based violence / sexual harassment? __________

• 84) How many approximate individual people have you spoken to about gender-based violence / sexual harassment? __________

• 85) How many individuals in your university residence hall have you spoken to about gender-based violence / sexual harassment? ______________

86) Please read the following list of statements and select to what extent you feel you disagree or agree. [1- Completely disagree, 2- Moderately disagree, 3- Slightly disagree, 4- Slightly agree, 5- Moderately agree, 6- Completely agree]

• 86) I have influence in decisions taken by my student accommodation and RA.
• 87) My student accommodation peers and RA openly share information with me.
• 88) I am typically involved and invited to actively participate in activities by my student accommodation and RA.
• 89) I am able to influence decisions that affect my degree or course.
• 90) I am usually among the last to know about important changes in the degree or course.
• 91) I am usually invited to important meetings in my degree or course.
• 92) My personal tutor often asks for my opinion before making important decisions.
• 93) My personal tutor does not share information with me.
• 94) I am invited to actively participate in meetings with my personal tutor.
• 95) I am often invited to contribute my opinion in meetings with senior university staff.
• 96) I frequently receive communication from university administrators higher than my senior university staff.
• 97) I am often invited to participate in meetings with senior university staff.
• 98) I am often asked to contribute in planning social activities not directly related to my immediate friend group.
• 99) I am always informed about informal social activities and events in my immediate friend group.
• 100) I am often invited to join my immediate friend group when they go out for a meal or drinks.

B.5: Student Accommodation Gender Dynamics Focus Group Activity

Messaging Questionnaire
For this activity, imagine you have a spare bedroom in your flat that you are trying to find a flatmate for. For each scenario below, select which one of the two inquiry messages you would respond to.
1. What is your participant ID #: ______________

2. Scenario 1
   a. Hi, I’m responding to your ad on [group name redacted] about the room. I’m an early 20’s female currently a student at [name redacted]. I’d love to come have a look at the place and meet you. Please let me know if the room is still available! Thank you – Olivia (she/her)
   b. Hey, I saw your post on the [group name redacted]. Is the room still available? I’m an early 20’s male student at [name redacted]. Is it possible to see the place and meet you, please let me know. Thanks – Leo (he/him)

3. Scenario 2
   a. Hi! I’m an early 20’s student at [name redacted], my pronouns are they/their and I’m very interested in your room posted on [group name redacted] if it is still available. Please let me know if you’d be open to meeting and showing me the room. Thanks! – Maisie
   b. Hello! Is your room on [group name redacted] still available for rent? I’m an early 20’s student (he/him pronouns). If it’s possible for us to meet and for me to see the room, please be in touch. Thank you! – Blair

4. Scenario 3
   a. Hi there. Could you please let me know if your open room on [group name redacted] is still available? I’m a male, I use ‘he/him’ pronouns, and I am a student at [name redacted]. It would be great to meet and see the place. Just let me know. Thank you! – Alfie
   b. Hey there. Is the room you posted on [group name redacted] still open? I’m a non-binary student (they/their pronouns) at [name redacted]. Please message me if you’d like to meet and I can see the place. That’d be great. Thanks! – Greer

B.6: Final Participant Numbers per Survey and Focus Group for Gender-Based Violence Study

<table>
<thead>
<tr>
<th>Survey Group</th>
<th>Cohort 1 Participants</th>
<th>Cohort 2 Participants</th>
<th>Total # of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Training Survey (1-week prior to training)</td>
<td>11</td>
<td>21</td>
<td>32</td>
</tr>
<tr>
<td>Immediately Post-Training Survey (1-week within training completion)</td>
<td>0</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>3-weeks Post-Training Survey</td>
<td>8</td>
<td>5</td>
<td>13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Focus Group (or Survey Equivalent)</th>
<th>Cohort 1 Participants</th>
<th>Cohort 2 Participants</th>
<th>Total # of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Training Focus Group (Immediately prior to training)</td>
<td>0</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Immediately Post-Training Focus Group (1-week within training completion)</td>
<td>9</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>3-weeks Post-Training Focus Group</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>
Appendix C: Inclusion Study

C.1: Inclusion Study Communications

The study was communicated to participants differently depending on whether the participant was a student or a tutor. This process of communicating consent was modelled after another tutorial study out of the same university. For tutors, I mentioned the study at each training session in both schools. Training participants were told that if they were part of certain courses, they would be contacted to express interest in participating in the study. Prior to the term, potential tutor participants were then emailed by the course organisers who expressed why the study was being done and their support for the study. Tutors were told to contact me if they were interested in participating or if they had any questions. I collected consent forms mostly digitally for tutors, unless there was a technological issue. Students first heard about the study during their first tutorial where the tutors who had already consented introduced the study and a short video of me. The 2-minute video explained why the study was being done, and what it would involve from the students. Tutors then handed out consent forms for students to sign. I collected the forms at the end of the first tutorial, only coming into the class after the students left. Students were then contacted via email about the study and reminders were sent to sign-up if they were interested. This continued for the first three weeks of tutorials given student schedules were shifting more than expected.

C.2: Inclusion Study Consent Forms

We would like to invite you to take part in a research study. Before you decide whether you want to take part, you should understand why the research is being done, and what it would involve for you. Please take time to read the following information carefully and contact us if you need further information before deciding whether you wish to participate.

What is the purpose of the study?
The aim of this comparative study is to evaluate the [school name redacted] mandatory inclusive classrooms training for tutors, in comparison to the [school name redacted] equality, diversity, and inclusion training for TA’s. This study hopes to establish and identify factors which contribute to inclusion trainings being effective, and to consider what additional resources will be needed to upskill tutors, to enhance equality, diversity, and inclusion within the university’s tutorials.
Why have I been asked to take part?
You have been asked to take part because you are a student in one of the tutorials for which the course organiser has agreed to allow tutors and students to participate in the study.

If I agree to participate in the study, what will I take part in?
- **3 Observational Tutorials (~2.5 hours total):** Your tutorial will be observed and potentially recorded if all of your fellow peers and tutor fully consent. Recordings will be transcribed, and then the recording would be destroyed. You will not be asked to attend any additional tutorials. Rather, the PhD researcher on this study will sit in on three of your pre-scheduled tutorials of which you are already scheduled to attend. No one beyond the PhD researcher will have access to the tutorial recordings.
- **2 Surveys (~40 minutes total):** You will be asked to complete two surveys focusing on your knowledge, behaviours, and feelings of inclusion, safety, and trust in your tutorial. The first survey will also contain demographic questions and should take about ~15-20 minutes. The following survey will take no longer than ~7-15 minutes to complete.

What are the possible benefits and disadvantages of taking part?
We expect that results from the study will help us to evaluate both inclusive classrooms trainings, as well as guide future diversity, equality, and inclusion strategy. If you participate in this study, you will contribute some of your personal time to the completion of the surveys. Additionally, discussing issues relating to diversity, inclusion and your own experiences as a student with these topics can be triggering and sensitive; we will try to ensure students feel supported throughout the study. We will remind students throughout the study of university support and resources that you can access. Your insights into your experience of tutorials will help foster a meaningful feedback loop for the schools to understand how to improve your tutorial experiences.

Do I have to take part?
It is entirely up to you whether you decide to participate in the research or not. If you agree to participate, you are free to withdraw at any time and without giving a reason. If you participate in the surveys and observational pieces of the survey, but then decide you want to withdraw, your data (responses to surveys) will be deleted as long as you let the PhD researcher know before the data have been analysed/written up. Deciding not to take part, or changing your mind at a later stage will not affect the education or support that you receive now or in the future.
Why were the trainings developed?
The [school name redacted] training was developed to achieve the following outcomes:

1. To increase awareness and knowledge for tutors of what constitutes equality, diversity, and inclusion
2. To teach tutors how they can alter behaviours to generate a safe and inclusive community
3. To act as a pilot training

The [school name redacted] training was developed by the Director of Equality and Diversity.

What happens to my information?
Your data will be processed in accordance with Data Protection law. All personal information you provide during the course of the research study is kept confidential. While the tutorials will be recorded (if all students consent), transcribed (typed-up), and analysed, any direct identifiers or information that might identify you will be removed (ex. Year in program, department, educational background, etc.). Participants will be given numbers. While names will be retained initially to help link the surveys to the organisational network analysis, once the data collection is complete, names will be completely replaced by unidentifiable numbers. All data with names will be destroyed. Indirect identifiers such as demographic information will be retained, but will not be shared at an individual-level. Rather, demographics will be used to extract high-level programme-wide analysis based on gender, race, age, etc. We will follow ethical and legal guidelines and keep all information privately and securely. All documents will be stored electronically in a secure [name redacted] space. Anonymised data will be stored for a maximum of 5 years. While data is stored, it will be retained on a password protected server only accessible by the PhD researcher.

What will happen to the results of the study?
The results of the study will be used to guide future diversity, equity, and inclusion trainings for [name redacted] students and staff. The results will be published in a doctoral thesis, as well as in diversity and inclusion, or educational journals. You will not be identifiable in any published results, and the PhD researcher will make the papers available to participants after publication.

How do I take part?
If you are interested in taking part, please sign the following consent form and either hand in, or send to [name redacted] at [email redacted].
Complaints
If you have any concerns or complaints about this research, please contact [name redacted] at [email redacted].

Data Protection
For general information about how we use your data please go to: [link redacted]
Alternatively, please contact the [name redacted]: [link redacted]

Many thanks for your time and help

Inclusive Classrooms Training Evaluation Research Study
Consent Form for Observations, Recordings & Surveys – Students

Please initial each box to confirm that you agree with the statements below. The statements will be re-read to you ahead of data collection and you will be asked to state your agreement.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Please initial box</th>
</tr>
</thead>
<tbody>
<tr>
<td>I confirm that I have read and understand the participant information sheet for the above study.</td>
<td></td>
</tr>
<tr>
<td>I confirm that I have had the opportunity to consider the information, ask questions and have had these questions answered satisfactorily.</td>
<td></td>
</tr>
<tr>
<td>I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason and without any care or support and/or legal rights being affected.</td>
<td></td>
</tr>
<tr>
<td>I understand that data collected about me during the study will be converted to anonymised data and a number will be used to protect my identity. Any direct identifiers will not be shared, while my demographics will only be used for high-level programme-wide analysis.</td>
<td></td>
</tr>
<tr>
<td>I agree to the tutorials I am in being audio and video recorded, knowing that the recordings will be destroyed once transcribed. These recordings will only be viewable by the PhD researcher on the study who will also conduct the observations.</td>
<td></td>
</tr>
<tr>
<td>Please initial box</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>---</td>
</tr>
<tr>
<td>I agree to my anonymised data being used in future studies.</td>
<td></td>
</tr>
<tr>
<td>I understand that all data will be stored in accordance with the [policy name redacted].</td>
<td></td>
</tr>
<tr>
<td>I agree to participate in the above study.</td>
<td></td>
</tr>
</tbody>
</table>

**FULL NAME:** __________________________

**DATE:** __________________

---

**Inclusive Classrooms Training Evaluation Research Study**

Consent Form for Recordings & Surveys ONLY – Students

Please initial each box to confirm that you agree with the statements below. The statements will be re-read to you ahead of data collection and you will be asked to state your agreement.

<table>
<thead>
<tr>
<th>Please initial box</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I confirm that I have read and understand the participant information sheet for the above study.</td>
<td></td>
</tr>
<tr>
<td>I confirm that I have had the opportunity to consider the information, ask questions and have had these questions answered satisfactorily.</td>
<td></td>
</tr>
<tr>
<td>I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason and without any care or support and/or legal rights being affected.</td>
<td></td>
</tr>
<tr>
<td>I understand that data collected about me during the study will be converted to anonymised data and a number will be used to protect my identity. Any direct identifiers will not be shared, while my</td>
<td></td>
</tr>
</tbody>
</table>
Inclusive Classrooms Training Evaluation Research Study
Participant Information Sheet – Tutors

We would like to invite you to take part in a research study. Before you decide whether you want to take part, you should understand why the research is being done, and what it would involve for you. Please take time to read the following information carefully and contact us if you need further information before deciding whether you wish to participate.

**What is the purpose of the study?**
The aim of this comparative study is to evaluate the [school name redacted] mandatory inclusive classrooms training for tutors, in comparison to the [school name redacted] equality, diversity, and inclusion training for TAs. This study hopes to establish and identify factors which contribute to inclusion trainings being effective, and to consider what additional resources will be needed to upskill tutors, to enhance equality, diversity, and inclusion within the university’s tutorials. Furthermore, this study hopes to establish
how computational social science methodologies can be leveraged by the university to quantifying inclusion, safety, and trust within the student experience.

Why was the training developed? This training was co-developed by a working group consisting of four tutors, [name redacted], and with guidance from the school’s [name redacted] to achieve the following outcomes:

1. To increase awareness and knowledge for tutors of what constitutes equality, diversity, and inclusion at the [name redacted], as well as within the classroom
2. To teach tutors how they can alter behaviours to generate a safe and inclusive community in which all students, regardless of identity, are given the tools & resources to be successful in their courses and larger academic programmes
3. To act as a pilot training with the potential to implement similar tutor trainings at other [name redacted] schools

[Name redacted] training was developed by the [name redacted] for TA’s as well as all staff.

Why have I been asked to take part? You have been asked to take part because you are a tutor for one of the designated [name redacted] courses whose course organiser has agreed to cascade this study.

If I agree to participate in the study, what will I take part in?

- 3 Observational Tutorials (~2.5 hours total): Your tutorial will be observed and potentially recorded (if you and all of your students fully consent). Recordings will be transcribed, and then the recording would be destroyed. You will not be asked to tutor any additional tutorials. Rather, the PhD researcher on this study will sit in on three of your pre-scheduled tutorials of which you are already facilitating. No one beyond the PhD researcher will have access to the tutorial recordings.

- 2 Surveys (~40 minutes total): You will be asked to complete two surveys focusing on your knowledge, behaviours, and feelings of inclusion, safety, and trust in your tutorial. The first survey will also contain demographic questions and should take about ~15-20 minutes. The following survey will take no longer than ~7-15 minutes to complete.

What are the possible benefits and disadvantages of taking part? We expect that results from the study will help us to evaluate [name redacted] trainings, as well as guide future diversity, equality, and inclusion strategy. If you participate in this study, you will contribute some of your personal time to the completion of the surveys. Additionally, discussing issues relating to diversity, inclusion and your own teaching can
be triggering and sensitive; we will try to ensure tutor participants feel supported throughout the study. We will remind tutors throughout the study of university support and resources that you can access. Additionally, we hope to share some of the high-level insights on classroom inclusion with participants once the analysis write-up is complete. This will, in turn, help foster a meaningful feedback loop for tutors to understand how their behaviours influence classroom inclusion, safety, and trust.

Do I have to take part?
It is entirely up to you whether you decide to participate in the research or not. If you agree to participate, you are free to withdraw at any time and without giving a reason. If you participate in the surveys and observational pieces of the survey, but then decide you want to withdraw, your data (responses to surveys) will be deleted as long as you let the PhD researcher know before the data have been analysed/written up. Deciding not to take part, or changing your mind at a later stage will not affect the education or support that you receive now or in the future.

What happens to my information?
Your data will be processed in accordance with Data Protection law. All personal information you provide during the course of the research study is kept confidential. While the tutorials will be recorded (if all students consent), transcribed (typed-up), and analysed, any direct identifiers or information that might identify you will be removed (ex. Year in program, department, educational background, etc.). Participants will be given numbers. While names will be retained initially to help link the surveys to the organisational network analysis, once the data collection is complete, names will be completely replaced by unidentifiable numbers. All data with names will be destroyed. Indirect identifiers such as demographic information will be retained, but will not be shared at an individual-level. Rather, demographics will be used to extract high-level programme-wide analysis based on gender, race, age, etc. We will follow ethical and legal guidelines and keep all information privately and securely. All documents will be stored electronically in a secure [name redacted] space. Anonymised data will be stored for a maximum of 5 years. While data is stored, it will be retained on a password protected server only accessible by the PhD researcher.

What will happen to the results of the study?
The results of the study will be used to guide future diversity, equity, and inclusion trainings for [name redacted] students and staff. The results will be published in a doctoral thesis, as well as in diversity and inclusion, or educational journals. You will not be identifiable in any published results, and the PhD researcher will make the papers available to participants after publication.
How do I take part?
If you are interested in taking part, please sign the following consent form and either hand in, or send to [name redacted] at [email redacted].

Complaints
If you have any concerns or complaints about this research, please contact the [name redacted] at [email redacted].

Data Protection
For general information about how we use your data please go to: [link redacted]
Alternatively, please contact the [name redacted]: [link redacted]

Many thanks for your time and help

Inclusive Classrooms Training Evaluation Research Study
Consent Form for Recordings, Observations & Surveys – Tutors

Please initial each box to confirm that you agree with the statements below. The statements will be re-read to you ahead of data collection and you will be asked to state your agreement.

<table>
<thead>
<tr>
<th>Please initial box</th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I confirm that I have read and understand the participant information sheet for the above study.</td>
</tr>
<tr>
<td></td>
<td>I confirm that I have had the opportunity to consider the information, ask questions and have had these questions answered satisfactorily.</td>
</tr>
<tr>
<td></td>
<td>I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason and without any care or support and/or legal rights being affected.</td>
</tr>
<tr>
<td></td>
<td>I understand that data collected about me during the study will be converted to anonymised data and a number will be used to protect my identity. Any direct identifiers will not be shared, while my demographics will only be used for high-level programme-wide analysis.</td>
</tr>
</tbody>
</table>
I agree to the tutorials I am in being audio and video recorded, knowing that the recordings will be destroyed once transcribed. These recordings will only be viewable by the PhD researcher on the study who will also conduct the observations.  

I agree to my anonymised data being used in future studies.  

I understand that all data will be stored in accordance with the [policy name redacted].  

I agree to participate in the above study.  

<table>
<thead>
<tr>
<th>Please initial box</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
</tr>
<tr>
<td>☐</td>
</tr>
<tr>
<td>☐</td>
</tr>
<tr>
<td>☐</td>
</tr>
</tbody>
</table>

---

Inclusive Classrooms Training Evaluation Research Study  
Consent Form for Recordings & Surveys ONLY – Tutors

Please initial each box to confirm that you agree with the statements below. The statements will be re-read to you ahead of data collection and you will be asked to state your agreement.

<table>
<thead>
<tr>
<th>Please initial box</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
</tr>
<tr>
<td>☐</td>
</tr>
<tr>
<td>☐</td>
</tr>
<tr>
<td>☐</td>
</tr>
</tbody>
</table>

I confirm that I have read and understand the participant information sheet for the above study.  

I confirm that I have had the opportunity to consider the information, ask questions and have had these questions answered satisfactorily.  

I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason and without any care or support and/or legal rights being affected.  

I understand that data collected about me during the study will be converted to anonymised data and a number will be used to protect my identity. Any direct identifiers will not be shared, while my
Thank you for taking this survey and participating in our research. This first survey should take about ~15-20 minutes to complete in one-sitting. The survey later in this study (Week #11) will require closer to ~7-15 minutes of your time as it will not include demographic questions. As a reminder, these surveys are critical to the process of evaluating the [name redacted] diversity, equality, and inclusion strategy. Trigger warning: Discussing issues relating to diversity, equity, and inclusion can be triggering and highly sensitive, thus if you need to take a break between questions to collect yourself, we encourage you to do so. We also encourage you to utilise university support and resources, which you will find links to at the end of this survey.

Your data will be processed in accordance with Data Protection law. All personal information you provide during the course of the research study is kept confidential. Indirect identifiers such as demographic information will be retained, but will not be shared at an individual-level. Rather, demographics will be used to extract high-level programme-wide analysis based on gender, race, age, etc. We will follow ethical and legal guidelines and keep all information privately and securely.

1) What is your student number? __________
2) What is your age?: ________________________________
3) What is your sex?: Male, female, non-binary, trans male, trans female, information refused, I do not identify as any of these and instead identify as: __________
4) What is your sexual orientation?: Heterosexual, homosexual, bisexual, pansexual, asexual, lesbian, gay, queer, information refused, I do not identify as any of these and instead identify as ________________
5) Which year of your undergraduate studies are you in?
6) Which school are you a part of?: [Drop-down options]
7) Which programme are you in?: ________________________________________
8) Are you a member of a university-sponsored activity? (ex. Societies, academic competitions, sporting teams, university bands, student government, etc.): Yes, no
9) Do you identify as having a disability and/or impairment?: Yes, no, information refused
10) What is your race and ethnicity? Please note, you are able to select multiple answers here or write-in your own race & ethnicity if you do not see your’s below.: White, White-Scottish, White- Other British, Black or Black British- Caribbean, Black or Black British-African, Other Black Background: ____, Asian or Asian British- Indian, Asian or Asian British- Pakistani, Asian or Asian British- Bangladeshi, Chinese, Other Asian Background: _______, Mixed- White and Black Caribbean, Mixed- White and Black African, Mixed- White and Asian, Latina/Latinx/Latino, Arab, Gypsy or Traveller, Not known, Information Refused, Other racial & ethnic background: ________
11) Please read the following list of statements and select to what extent you feel you disagree or agree. Note that each question pertains to either your [class name redacted] tutorial group (depending on which course you are enrolled in): [1- Completely disagree, 2- Somewhat disagree, 3- Neither disagree nor agree, 4- Somewhat agree, 5- Completely agree]
   • 11) I feel a sense of community with the people in my tutorial group, including my tutor.
   • 12) It is very important to me to feel a sense of community with the people in my tutorial group, including my tutor.
   • 13) I feel a sense of trust with the people in my tutorial group, including my tutor.
   • 14) It is important to me to feel a sense of trust with the people in my tutorial group, including my tutor.
   • 15) I feel a sense of safety with the people in my tutorial group, including my tutor.
   • 16) It is important to me to feel a sense of safety with the people in my tutorial group, including my tutor.
   • 17) I feel a sense of community with people at my university.
   • 18) It is very important to me to feel a sense of community with people at my university.
   • 19) I feel a sense of trust with people at my university.
   • 20) It is important to me to feel a sense of trust with people at my university.
   • 21) I feel a sense of safety with people at my university.
22) It is important to me to feel a sense of safety with people at my university.

23) Please read the following list of statements and select to what extent you feel you disagree or agree. Please note, the scale is now from 1-6, as follows: [1- Completely disagree, 2- Moderately disagree, 3- Slightly disagree, 4- Slightly agree, 5- Moderately agree, 6- Strongly agree]

- 23) My peers in my tutorial group openly share information with me.
- 24) I am typically involved and invited to actively participate in activities in my tutorial group by my peers.
- 25) I am always informed about informal social activities and events with my peers in my tutorial group.
- 26) I am often invited to join my peers from my tutorial group when they go out for a meal, drinks, or a study session.
- 27) I am often invited to contribute my opinion in tutorials by my tutor.
- 28) I am usually among the last to know about important changes in the tutorial.
- 29) My tutor often asks for my opinion before making important decisions.
- 30) I frequently receive communication from my tutor.
- 31) I am invited to actively participate in communication with my tutor.
- 32) I am able to influence decisions that affect my tutorial.
- 33) When there is a meeting with the course organisers outside of my tutorial, I am invited to participate by my tutor.

Tutor Survey

Thank you for taking this survey and participating in this research. This first survey should take about ~15-20 minutes to complete in one-sitting. The survey later in this study (Week #11) will require closer to ~7-15 minutes of your time as it will not include demographic questions. As a reminder, these surveys are critical to the process of evaluating the [name redacted] diversity, equality, and inclusion strategy.

Trigger warning: Discussing issues relating to diversity, equality, and inclusion can be triggering and highly sensitive, thus if you need to take a break between questions to collect yourself, we encourage you to do so. We also encourage you to utilise university support and resources, which you will find links to at the end of this survey.

Your data will be processed in accordance with Data Protection law. All personal information you provide during the course of the research study is kept confidential. Indirect identifiers such as demographic information will be retained, but will not be shared at an individual-level. Rather, demographics will be used to extract high-level

---

7 The second iteration of this survey was the same, just minus the demographic questions.
programme-wide analysis based on gender, race, age, etc. We will follow ethical and legal guidelines and keep all information privately and securely.

1) What is your name? __________
2) What is your age?: ________________________________
3) What is your sex?: Male, female, non-binary, trans male, trans female, information refused, I do not identify as any of these and instead identify as: __________
4) What is your sexual orientation?: Heterosexual, homosexual, bisexual, pansexual, asexual, lesbian, gay, queer, information refused, I do not identify as any of these and instead identify as __________
5) Which year of your postgraduate studies are you in, or how many years after postgraduate studies are you?: _________________
6) How many academic terms have you tutored for? (If this is your first term tutoring, put 1): __________
7) Which school are you a part of (if applicable): [drop down options]
8) Which programme are you in (if applicable): [drop down options]
9) Do you identify as having a disability and/or impairment?: Yes, no, information refused
10) What is your race and ethnicity? Please note, you are able to select multiple answers here or write-in your own race & ethnicity if you do not see your’s below.: White, White-Scottish, White- Other British, Black or Black British- Caribbean, Black or Black British- African, Other Black Background: _____, Asian or Asian British- Indian, Asian or Asian British- Pakistani, Asian or Asian British- Bangladeshi, Chinese, Other Asian Background: ______, Mixed- White and Black Caribbean, Mixed- White and Black African, Mixed- White and Asian, Latina/Latinx/Latino, Arab, Gypsy or Traveller, Not known, Information Refused, Other racial & ethnic background: ______
11) Please read the following list of statements and select to what extent you feel you disagree or agree. Please note, if you have multiple tutorial groups, answer the questions below as it pertains to the one tutorial group that is being observed or that you’ve been asked to focus on: [1- Completely disagree, 2- Somewhat disagree, 3- Neither disagree nor agree, 4- Somewhat agree, 5- Completely agree]
   • 11) I feel a sense of community with the people in my tutorial group.
   • 12) It is very important to me to feel a sense of community with the people in my tutorial group.
   • 13) I feel a sense of trust with the people in my tutorial group.
   • 14) It is important to me to feel a sense of trust with the people in my tutorial group.
   • 15) I feel a sense of safety with the people in my tutorial group.
   • 16) It is important to me to feel a sense of safety with the people in my tutorial group.
   • 17) I feel a sense of community with people at my university.
   • 18) It is very important to me to feel a sense of community with people at my university.
   • 19) I feel a sense of trust with people at my university.
• 20) It is important to me to feel a sense of trust with people at my university.
• 21) I feel a sense of safety with people at my university.
• 22) It is important to me to feel a sense of safety with people at my university.

23) Please read the following list of statements and select to what extent you feel you disagree or agree. Please note, the scale is now from 1-6, as follows: [1- Completely disagree, 2- Moderately disagree, 3- Slightly disagree, 4- Slightly agree, 5- Moderately agree, 6- Strongly agree]

• 23) I often invite the students in my tutorial group to contribute their opinion in tutorials.
• 24) I ensure all students in my tutorial group know about important changes in the tutorial.
• 25) I often ask for the students in my tutorial group's opinion before making important decisions.
• 26) I frequently communicate to the students in my tutorial group. (ex. Learn, emails, etc.)
• 27) I invite students in my tutorial group to actively participate in communication with me.
• 28) I ensure students in my tutorial group are able to influence decisions that affect their tutorial.
• 29) When there is a meeting with the course organisers outside my tutorial group, I invite the students in my group to participate.
C.4: Inclusion Study Survey Completion Rates

Table C.4: Total Student Survey Completion Rates (%) per Survey and Tutorial Group

<table>
<thead>
<tr>
<th>Tutorial</th>
<th>Number of Student Participants</th>
<th>Survey #1</th>
<th>Survey #2 (Δ = Difference from Survey #1 to Survey #2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>14</td>
<td>57.14%</td>
<td>28.57% (Δ = -28.57%)</td>
</tr>
<tr>
<td>B</td>
<td>13</td>
<td>69.23%</td>
<td>69.23% (Δ = 0.00%)</td>
</tr>
<tr>
<td>C</td>
<td>11</td>
<td>72.73%</td>
<td>63.64% (Δ = -9.09%)</td>
</tr>
<tr>
<td>D</td>
<td>9</td>
<td>44.44%</td>
<td>33.33% (Δ = -11.11%)</td>
</tr>
<tr>
<td>E</td>
<td>10</td>
<td>50.00%</td>
<td>40.00% (Δ = -10.00%)</td>
</tr>
<tr>
<td>F</td>
<td>9</td>
<td>66.67%</td>
<td>22.22% (Δ = -44.45%)</td>
</tr>
<tr>
<td>G</td>
<td>7</td>
<td>57.14%</td>
<td>57.14% (Δ = +0.00%)</td>
</tr>
<tr>
<td>H</td>
<td>11</td>
<td>63.64%</td>
<td>18.18% (Δ = -45.46%)</td>
</tr>
<tr>
<td>I</td>
<td>10</td>
<td>60.00%</td>
<td>70.00% (Δ = +10.00%)</td>
</tr>
<tr>
<td>J</td>
<td>8</td>
<td>62.50%</td>
<td>37.50% (Δ = -25.00%)</td>
</tr>
<tr>
<td>K</td>
<td>8</td>
<td>50.00%</td>
<td>50.00% (Δ = 0.00%)</td>
</tr>
<tr>
<td>L</td>
<td>8</td>
<td>62.50%</td>
<td>50.00% (Δ = -12.50%)</td>
</tr>
<tr>
<td>M</td>
<td>9</td>
<td>77.78%</td>
<td>33.33% (Δ = -44.45%)</td>
</tr>
<tr>
<td>N</td>
<td>11</td>
<td>72.73%</td>
<td>72.73% (Δ = 0.00%)</td>
</tr>
<tr>
<td>O</td>
<td>10</td>
<td>70.00%</td>
<td>40.00% (Δ = -30.00%)</td>
</tr>
<tr>
<td>P</td>
<td>7</td>
<td>100.00%</td>
<td>85.71% (Δ = -14.29%)</td>
</tr>
<tr>
<td>Q</td>
<td>8</td>
<td>50.00%</td>
<td>75.00% (Δ = +25.00%)</td>
</tr>
<tr>
<td>R</td>
<td>8</td>
<td>50.00%</td>
<td>62.50% (Δ = +12.50%)</td>
</tr>
<tr>
<td>S</td>
<td>11</td>
<td>63.64%</td>
<td>63.64% (Δ = 0.00%)</td>
</tr>
<tr>
<td>T</td>
<td>13</td>
<td>53.85%</td>
<td>61.54% (Δ = +7.69%)</td>
</tr>
<tr>
<td>U</td>
<td>12</td>
<td>50.00%</td>
<td>8.33% (Δ = -41.67%)</td>
</tr>
<tr>
<td>V</td>
<td>4</td>
<td>25.00%</td>
<td>75.00% (Δ = +50.00%)</td>
</tr>
<tr>
<td>W</td>
<td>1</td>
<td>100.00%</td>
<td>100.00% (Δ = 0.00%)</td>
</tr>
<tr>
<td>X</td>
<td>10</td>
<td>60.00%</td>
<td>50.00% (Δ = -10.00%)</td>
</tr>
<tr>
<td>Y</td>
<td>8</td>
<td>50.00%</td>
<td>25.00% (Δ = -25.00%)</td>
</tr>
<tr>
<td>Z</td>
<td>9</td>
<td>88.89%</td>
<td>33.33% (Δ = -55.56%)</td>
</tr>
<tr>
<td>ZZ</td>
<td>10</td>
<td>30.00%</td>
<td>10.00% (Δ = -20.00%)</td>
</tr>
</tbody>
</table>
C.5: Inclusion Study Student Responses per Survey and Survey Item

Figure C.5a: Likert Scale Responses per Survey for Survey Items #1-#4

- **Item #1:** My peers in my tutorial group openly share information with me.

- **Item #2:** I am typically involved and invited to actively participate in activities in my tutorial group by my peers.

- **Item #3:** I am always informed about informal social activities and events with my peers in my tutorial group.

- **Item #4:** I am often invited to join my peers from my tutorial group when they go out for a meal, drinks, or a study session.
Figure C.5b: Likert Scale Responses per Survey for Survey Items #5-8
Figure C.5c: Likert Scale Responses per Survey for Survey Items #9-11
C.6: Inclusion Study Additional Networks

Figure C.6a: Mid-Term Network for Tutorial A (Guessed Demographics)

Figure C.6b: Mid-Term Network for Tutorial A (Non-Guessed Demographics)
Figure C.6c: Mid-Term Network for Tutorial B

Figure C.6d: Mid-Term Network for Tutorial C
Figure C.6e: Cumulative Network for Tutorial A (Non-Guessed Demographics)
References


https://scholar.google.com/scholar_url?url=https://files.osf.io/v1/resources/5eyur/providers/osfstorage/6330767e18f458113842986f%3Faction%3Ddownload%26direct%26version%3D3&hl=en&sa=X&ei=usolZZ9F5H7mQG8oKWgBQ&scisig=AFWwaeZzkXwBvWWr4ZUeyUnycjHn&oi=scholarr


Baez, B. (2002). Confidentiality in qualitative research: Reflections on secrets, power and agency. *Qualitative Research, 2*(1), 35–58.


258


Smith, L., & Powers, B. (n.d.). The Role of Trust in Determining the Effectiveness of Bystander Training Programs at Georgetown University. Georgetown University.


