

**Sexual violence and the gender pay gap**

A Corpus Assisted Discourse Studies (CADS) approach

The thesis is submitted for the degree of  
PhD Linguistics by Research Only

LAEL, Lancaster University

Author: Jillian Crawley (35025592)

Submitted: August 2025

## Abstract

At the start of this study two longstanding feminist issues were highly visible on Twitter<sup>1</sup>, sexual violence, expressing itself through the #MeToo movement, and the BBC pay gap scandal. I was interested in what their discourses had in common, and any possible cultural influences.

Using three search terms: *sexual harassment*; *#MeToo*; and *pay gap*, I harvested data from Twitter and Nexis between April 2019 and March 2021, resulting in six 2.4M word corpora. My focus was on the Twitter corpora with their Nexis equivalents providing a comparison. My analysis took a Corpus Assisted Discourse Studies approach combining: a corpus-driven element; Critical Discourse Analysis; and Feminist Post-structural Discourse Analysis. For the sexual harassment and #MeToo corpora I focused on the lemmas VICTIM and SURVIVOR, and broader modifiers within the pay gap corpora. My analysis was based on van Leeuwen's social actor and legitimation frameworks (2008, p. 108), plus manually coded flags for prosody and veracity.

There are three key findings. Firstly, tension exists between taking responsibility for ensuring personal safety and success, and blaming/shaming those who fail to do so, irrespective of structural barriers. Secondly, the majority of the disadvantaged may be missing from the discourse, silenced or muted. Thirdly, the meanings of VICTIM and SURVIVOR are changing and contested. In the west these linguistic changes are situated alongside rising neoliberalism, which favours both the individualistic self-reliance implied by SURVIVOR, and post-feminism which believes gender equality has been achieved. These corpora show that SURVIVOR is increasingly lauded for resilience and agency, but the benefits may come with obligations and expectations. Simultaneously, VICTIM becomes denigrated and weakened as a key tool for challenging disadvantage.

---

<sup>1</sup> In July 2022 the social media platform Twitter was rebranded "X". This was after the data gathering for this research was complete, so the decision was made to continue referring to the platform as "Twitter".

## Table of Contents

Abstract .....	ii
List of tables .....	ix
List of figures .....	xii
Acknowledgements.....	xiii
Author’s declaration .....	xiv
1 Introduction .....	1
1.1 The journey .....	3
1.2 How this study hopes to contribute.....	4
1.3 The research questions .....	7
1.4 Summary of chapters .....	8
1.4.1 Chapter 2: Literature review .....	8
1.4.2 Chapter 3: Methodology .....	10
1.4.3 Chapter 4: RQ1, Prosody and sexual violence .....	11
1.4.4 Chapter 5: RQ2, Prosody and the pay gap .....	11
1.4.5 Chapter 6: RQ3, Delegitimation .....	12
1.4.6 Chapter 7: RQ4, Labelling and the missing .....	12
1.4.7 Chapter 8: Conclusion .....	12
2 Literature review .....	13
2.1 Introduction .....	13
2.2 Situating the public debate .....	13
2.2.1 Sexual harassment/violence .....	14
2.2.2 The pay gap .....	16
2.2.3 Links between sexual violence and the pay gap .....	20
2.3 Methodological linguistic tools .....	21
2.3.1 CL and CADS .....	21
2.3.2 The feminist context and FPDA .....	23
2.4 The sociological landscape.....	25
2.4.1 Discourse is reality, even when it’s not real.....	25
2.4.2 Language and cultural change .....	26
2.4.3 Feminist Legal Theory .....	26
2.4.4 Just World Theory .....	29
2.4.5 Reward Expectation Theory .....	29
2.4.6 Labelling Theory .....	30

2.5	Victims, survivors and the silenced.....	31
2.5.1	Ideal Victim .....	32
2.5.2	VICTIM and the legal system.....	33
2.5.3	The victimised in news and social media texts .....	35
2.5.4	The evaluation and debasement of VICTIM.....	38
2.5.5	SURVIVOR as active coping .....	40
2.5.6	Forced heroism of the survivor .....	40
2.5.7	The missing and the muted victims .....	41
2.5.8	Which victims may speak, and who are believed .....	42
2.5.9	Recognising victimisation as crime .....	44
2.5.10	Linguistic hi-jacking of VICTIM.....	45
2.5.11	Depoliticising victimisation .....	45
2.5.12	Victimisation by hate speech .....	46
2.6	Conclusion.....	46
3	Methodology.....	48
3.1	Introduction .....	48
3.2	Initial approach .....	49
3.2.1	The Twitter and Nexis data .....	49
3.2.2	The raw data .....	53
3.2.3	The reference corpora.....	54
3.2.4	Who are represented in the sexual violence and pay gap corpora? .....	54
3.2.5	A feminist perspective.....	59
3.3	Data processing.....	59
3.3.1	Harvesting tweets (TAGS).....	60
3.3.2	The deduplication and cleaning process for tweets (Excel).....	61
3.3.3	Harvesting news articles (Nexis) .....	62
3.3.4	Deduplicating news articles (WordSmith) .....	63
3.3.5	Final data cleaning (MS Word).....	63
3.3.6	Aggregating the monthly data .....	64
3.3.7	Early view of the data .....	64
3.3.8	Corpus interrogation (SketchEngine and #LancsBox) .....	66
3.4	Concordance categorisation .....	69
3.4.1	Van Leeuwen social actor and authority legitimation types .....	71
3.4.2	Veracity and +/- prosody flags .....	76
3.5	Identifying the themes.....	77

3.5.1	Victims and survivors .....	78
3.5.2	Silence .....	82
3.5.3	Agency.....	84
3.6	Issues.....	85
3.7	Conclusion of Methodology.....	86
4	Prosody and sexual violence .....	87
4.1	Introduction .....	87
4.2	+/- prosody in representations .....	88
4.2.1	Sentiment analysis tools .....	88
4.2.2	Defining positive, negative and neutral prosody .....	88
4.2.3	Pilot of +/- prosody breakdown .....	90
4.2.4	Third-party cross check of +/- prosody coding .....	91
4.2.5	Prosody breakdown of vL500 datasets .....	92
4.2.6	Excluded concordances.....	93
4.3	Social actor representations .....	94
4.3.1	Coding and re-coding the vLNSHPilot datasets.....	94
4.3.2	Third-party cross-check of van Leeuwen social actor types .....	97
4.3.3	Frequency of van Leeuwen Social Actor types .....	98
4.4	Positive representations of VICTIM and SURVIVOR concordances.....	98
4.4.1	Activation .....	99
4.4.2	Appraisalment.....	110
4.4.3	Indifferentiation .....	111
4.4.4	Passivations.....	112
4.5	Negative representations of VICTIM and SURVIVOR concordances.....	114
4.5.1	Activation .....	114
4.5.2	Appraisalment.....	119
4.5.3	Indifferentiation .....	123
4.5.4	Passivations.....	124
4.6	Agency of SURVIVOR as a double-edged sword.....	126
4.6.1	If not a victim, then what?.....	127
4.6.2	“Speaking as a victim/survivor” .....	130
4.7	Conclusion.....	133
5	Prosody and the pay gap.....	135
5.1	Introduction .....	135
5.2	+/- prosody representations .....	135

## Sexual violence and the gender pay gap

5.3	Van Leeuwen social actor representations .....	136
5.4	Positive representations within pay gap datasets .....	137
5.4.1	Activation .....	137
5.4.2	Appraisalment.....	144
5.4.3	Indifferentiation .....	146
5.4.4	Passivations .....	147
5.4.5	Suppression.....	155
5.5	Negative representations within pay gap datasets.....	159
5.5.1	Activation .....	160
5.5.2	Appraisalment.....	162
5.5.3	Passivations .....	163
5.5.4	Suppression.....	167
5.6	Humour and sarcasm .....	169
5.7	Conclusion.....	171
6	Delegitimation and the pay gap .....	174
6.1	Introduction .....	174
6.2	Approaching the pay gap corpora.....	174
6.3	Differences between the Nexis and Twitter pay gap corpora.....	175
6.4	Veracity .....	177
6.4.1	<i>Myth</i> - TwitterPG.....	177
6.4.2	<i>Real</i> - TwitterPG .....	179
6.4.3	<i>Real</i> - NexisPG .....	182
6.4.4	<i>Lie</i> - TwitterPG.....	183
6.4.5	<i>Fake</i> - TwitterPG .....	184
6.4.6	<i>False</i> - TwitterPG .....	185
6.4.7	<i>True</i> - TwitterPG .....	186
6.4.8	<i>Hoax</i> - TwitterPG .....	188
6.5	Measure of Acceptability .....	189
6.5.1	<i>Bad</i> - TwitterPG .....	189
6.5.2	<i>Bad</i> - NexisPG .....	190
6.5.3	<i>Wrong</i> - TwitterPG .....	190
6.5.4	<i>Important</i> - TwitterPG .....	191
6.5.5	<i>Fair</i> - TwitterPG .....	193
6.5.6	<i>Problem</i> - TwitterPG.....	194
6.6	Contempt .....	195

6.6.1	<i>Bullshit</i> - TwitterPG .....	196
6.6.2	<i>Nonsense</i> - TwitterPG .....	196
6.6.3	<i>Ridiculous</i> - TwitterPG .....	197
6.6.4	<i>Joke</i> - TwitterPG .....	198
6.7	Conclusion.....	199
7	The labelled and the missing .....	201
7.1	Introduction .....	201
7.2	Shifting definitions of VICTIM and SURVIVOR .....	201
7.3	Pros and cons of VICTIM.....	202
7.3.1	Stigmatisation of VICTIM.....	203
7.3.2	Access to support.....	204
7.3.3	Distortion of veracity .....	205
7.3.4	Personal responsibility as victim-blaming.....	206
7.3.5	Biased use of <i>alleged</i> .....	209
7.3.6	Linguistic hi-jacking of VICTIM.....	210
7.4	Pros and cons of SURVIVOR.....	212
7.4.1	SURVIVOR as the only alternative to VICTIM .....	212
7.5	Who is missing from pay gap discussions? .....	218
7.5.1	Corpora level .....	218
7.5.2	A process of elimination .....	219
7.5.3	Career segregation .....	222
7.6	Who is missing from sexual harassment discussions?.....	225
7.6.1	Who benefits from silence? .....	226
7.6.2	The dangers of speaking out.....	227
7.7	Conclusion.....	227
8	Conclusion.....	229
8.1	The key findings .....	229
8.1.1	RQ1.....	229
8.1.2	RQ2.....	231
8.1.3	RQ3.....	232
8.1.4	RQ4.....	232
8.2	Reflections.....	233
8.3	Limitations.....	235
8.4	Further research directions.....	236
8.5	Contribution and implications .....	237

## Sexual violence and the gender pay gap

8.5.1	Prioritising the range of voices .....	237
8.5.2	Survivor and victim, RQ1.....	238
8.5.3	Those wronged by pay gap RQ2.....	239
8.5.4	Pay gap and sexual violence RQ3, RQ4 .....	241
8.6	Concluding remarks .....	242
Appendix A – Rape myths .....		245
Appendix B – Corpus and dataset naming.....		249
Appendix C – Deduplication and cleaning formulae.....		251
Appendix D – Non-english language filtering formulae.....		252
Appendix E - Nouns modifying FEMALE .....		256
References.....		260

**List of tables**

Table 1, Lancaster University Library, 2025 search of peer reviewed journal .....	5
Table 2, Overlap in occurrences of SEXUAL HARASSMENT and #METOO .....	50
Table 3, Average number of words in tweets/articles, based on the _1Yr corpora.....	53
Table 4, Tweets in a randomised 100 that had a potentially identifiable @useraccount .....	54
Table 5, The most common targets of sexual violence or pay disparity .....	55
Table 6, Potential labels for the victimised .....	57
Table 7, Breakdown of social roles within concordances for I WAS.....	58
Table 8, Effects of Nexis similarity settings .....	62
Table 9, Effect of manipulating WordSmith length and difference.....	63
Table 10, I WAS pilot concordances .....	65
Table 11, Frequencies of VICTIM and SURVIVOR.....	66
Table 12, WordSketch for VICTIM within the TwitterSH_1Yr corpus.....	67
Table 13, SketchDifference for VICTIM and SURVIVOR.....	68
Table 14, Comparison of top 10 tokens for MI, LogDice, T-score and Log-likelihood.....	69
Table 15, Variables, categories, and research questions .....	70
Table 16, Frequencies of words used to dispute veracity in sexual harassment .....	77
Table 17, "PAY GAP is a..." collocates from a pay gap WordSketch.....	77
Table 18, Frequency of the lemmas VICTIM and SURVIVOR .....	78
Table 19, Verbs with VICTIM and SURVIVOR as object and subject in the _1Yr corpora .....	80
Table 20, Frequency in pmw of blame and SHAME in _1Yr corpora .....	81
Table 21, BLAME with victim as object in _1Yr corpora.....	81
Table 22, SHAME with victim as object in _1Yr corpora .....	82
Table 23, BLAME with survivor as subject in _1Yr corpora.....	82
Table 24, SHAME with survivor as subject in _1Yr corpora.....	82
Table 25, Verbs with VICTIM as object within sexual violence _1Yr corpora .....	83
Table 26, Top ten most frequent objects of the verb SILENCE .....	83
Table 27, Modifiers of the noun SILENCE within the sexual violence _1Yr corpora.....	84
Table 28, Loss of words upon aggregation of corpora .....	85
Table 29, +/- prosody of concordances in the vLPilot datasets .....	91
Table 30, +/- prosody of concordances 101-200 from the datasets .....	91
Table 31, +/- prosody coding of vL500 datasets .....	93
Table 32, Excluded (N/A) concordances.....	93

## Sexual violence and the gender pay gap

Table 33, Highest frequency social actor representations .....	96
Table 34, Highest frequency of positive representations .....	99
Table 35, Main themes for positive Activation representations.....	100
Table 36, Comparing Twitter positive Activation verbs that occur more than once .....	101
Table 37, Comparing Nexis positive Activation verbs that occur more than once .....	101
Table 38, Consequences of speaking out in Twitter Activation concordances .....	103
Table 39, Consequences of speaking out in Nexis Activation concordances.....	106
Table 40, Distribution of positive Passivations between VICTIM and SURVIVOR.....	113
Table 41, Frequency of negative representations of social actor types .....	114
Table 42, Dominant themes within the negative Activation concordances .....	115
Table 43, Collocation of PLAY with victim and survivor .....	118
Table 44, The most common evaluative adjectives .....	120
Table 45, Frequencies of negative Passivations for VICTIM and SURVIVOR .....	124
Table 46, Manually identified phrases Passivating negative VICTIM and SURVIVOR .....	125
Table 47, SketchEngine concordances for the KWIK NOT A VICTIM.....	127
Table 48, Frequency of AS A SURVIVOR and AS A VICTIM in SketchEngine .....	130
Table 49, The +/- prosody composition in the vLPG500 datasets .....	136
Table 50, Highest frequency social actor representations.....	136
Table 51, Frequency of positive coded social actor types .....	137
Table 52, Most common themes within positive Activation social actor types .....	138
Table 53, Top ten collocates of SCREAM within the enTenTen21 .....	139
Table 54, Most common themes for positive Appraisal social actor types .....	144
Table 55, Most common themes for positive Indifferentiation social actor types.....	146
Table 56, Most common themes for positive Passivations social actor types.....	147
Table 57, The most common themes for positive Suppression social actor types.....	156
Table 58, Frequency of negative representations of social actor types .....	159
Table 59, Frequency of negative themes within the vLTPG500.....	160
Table 60, Most common themes for negative Passivation social actor types .....	164
Table 61, Most common themes for negative Suppression social actor types .....	167
Table 62, Evaluative adjective predicates and “pay gap” is a ... for the lemma PAY GAP .....	175
Table 63, Statistic related modifiers of PAY GAP for NexisPG and TwitterPG .....	176
Table 64, Top ten collocates of CONTROLLED FOR and UNCONTROLLED .....	178
Table 65, Breakdown of adjective predicates of real .....	179

Table 66, Manual breakdown of the adjective predicate true.....	186
Table 67, Lemma SUPPORT and its collocation with victim and survivor.....	205
Table 68, Collocations of TRUTH in the sexual violence corpora, and enTenTen21.....	206
Table 69, Frequency of ACCUSE and ALLEGE, along with their top adjective modifier. ....	210
Table 70, Delegitimising modifiers of the lemma VICTIM.....	213
Table 71, Legitimising modifiers of the lemma SURVIVOR.....	214
Table 72, Modifiers victim and survivor of the lemma BRAVE.....	215
Table 73, Job title nouns modifying FEMALE (total in table 492).....	219
Table 74, Job title nouns modifying FEMALE (total in table 725).....	220
Table 75, Categories of profession nouns modifying FEMALE.....	221
Table 76, The coded corpora.....	249
Table 77, The coded datasets.....	250

## List of figures

Figure 1, Google Ngrams showing the changing use of terms .....	2
Figure 2, High level map of this dissertation .....	6
Figure 3, The sexual violence escalation pyramid.....	16
Figure 4, ONS gender pay gap 1997 - 2023.....	18
Figure 5, OECD gender wage gap in 2019 .....	19
Figure 6, US recessions from 1980 – 2024.....	39
Figure 7, Uma Thurman, the victim/survivor in the movie Kill Bill .....	41
Figure 8, BBC reporting of Al Fayed accusations.....	43
Figure 9, “Look out for him. He’s in a good mood.” .....	44
Figure 10, Methodology path .....	48
Figure 11, Active Twitter users worldwide, 1 <sup>st</sup> quarter 2010 to 1 <sup>st</sup> quarter 2019 .....	51
Figure 12, Comparing social media activities, February 2019 .....	52
Figure 13, 80% of tweets authored by 10% of users .....	60
Figure 14, LogDice calculations used in SketchEngine WordSketches .....	69
Figure 15, Re-drawing of van Leeuwen’s (2008) Social Actor Network .....	73
Figure 16, Tweet following USA women’s soccer win.....	142
Figure 17, Tweet “People want to work for less money” .....	144
Figure 18, Tweet “Blaming falling fertility on women ‘choosing’ education...” .....	150
Figure 19, Twitter graphic of MPs funded by Chinn donations .....	156
Figure 20, Tweet “Fallacy of Understated Evidence” graphic .....	165
Figure 21, Tweet Mocking Spongebob meme .....	170
Figure 22, Tweet “Feminism page 1” .....	171
Figure 23, Tweet linked from Example 206.....	187
Figure 24, Tweet “Women hold 2/3 <sup>rd</sup> of US student debt” .....	188
Figure 25, International Women’s Day Rreport .....	192
Figure 26, Alex Morgan .....	194
Figure 27, Tweet “No, no, no. Of course its not. It’s a systemic issue.” .....	195
Figure 28, Users in the United States predominate Twitter .....	207
Figure 29, Google’s Ngram Viewer for METoo .....	211
Figure 30, Gender pay gap for gross hourly earnings .....	223
Figure 31, Gender split of zero-hours contracts .....	224

### **Acknowledgements**

I am deeply indebted to Distinguished Professor Tony McEnery of Lancaster University's Department of Linguistics and English Language for his unfailing calm support and insights while steering me along this journey.

A special thank you is also due to Professor Paul Baker for his continued interest and advice.

I am also grateful to Dr Luke Collins and Dr Gavin Brookes for their kind and supportive comments.

Thanks are due to my fellow researcher Dr Usman Maravia for proofreading and being my sounding board throughout.

I would also like to thank the library staff for their prompt and efficient support through technical challenges.

I also need to thank my friend Deb Pain for her proof reading and moral support.

Lastly I am indebted to my late husband Leon Crawley for seeing in me what I could not see myself, starting me on this journey, and providing the inspiration to complete it.

**Author's declaration**

I Jillian Crawley declare that this thesis is my own work and has not been submitted in whole or in part for the award of a higher degree elsewhere.

## 1 Introduction

The present study is situated at a time when two major themes were constantly appearing in the news and social media, the BBC<sup>2</sup> pay gap<sup>3</sup> scandal, and a string of high profile sexual violence<sup>4</sup> cases such as those against Harvey Weinstein<sup>5</sup> and Brett Kavanaugh<sup>6</sup>. Any visitor to Twitter.com<sup>1</sup>(page ii) quickly found accusations of perpetrating or covering up sexual violence, leveled particularly against US celebrities or politicians by their opponents. In contrast, pay gap discussions featured fewer individuals, although a significant number were surprisingly emotionally charged. The coincidence of these two gendered narratives unfolding concurrently was intriguing.

Since the term *sexual harassment* was coined in the 1970s to describe unwanted sexual attention as a wrong, it has become increasingly discussed as unacceptable. As equal pay legislation took effect in several countries from the 1960s and 70s, the broader gender pay gap was becoming more visible by the 1990s, illustrated in Figure 1. While the Google Ngrams<sup>7</sup> indicate that sexual violence is discussed an order of magnitude more frequently than the gender pay gap, both were increasing when the #MeToo<sup>8</sup> hashtag went viral on Twitter in 2017, and the following year the BBC pay scandal broke.

When this study began in 2019, both issues were still regularly making headlines, prompting the question of what they might or might not have in common, and whether there was any relationship between them. Articles covering sexual violence in the news media almost always refer to workplace abuse, causing a few journalists to propose links with the gender pay gap. Further evidence emerged from the US where the US Equal Employment Opportunity Commission, who are responsible for enforcing civil rights, found that “significant power disparities” were a risk factor of harassment in the workplace (n.d., p. 7 (in pdf reader)).

---

<sup>2</sup> Carrie Grace was the BBC’s China correspondent up until 2018, when she resigned in protest at being paid less than her male peers (BBC, 2018a). This started a high profile equal pay scandal affecting the whole of the BBC.

<sup>3</sup> The pay gap is defined in section 2.2.2 section.

<sup>4</sup> Sexual violence is defined in 2.2.1.

<sup>5</sup> Harvey Weinstein was a Hollywood film producer accused of raping more than a dozen women, he was convicted on two counts in New York in 2020. In 2022 he was charge with 11 more counts of rape in California, and found guilty of three. In 2024 the New York convictions were overturned pending a retrial. Later in 2024 he was indicted on new charges that were consolidated into the retrial.

<sup>6</sup> In 2018 while being confirmed as a Supreme Court nominee, Brett Kavanaugh was accused of sexually assaulting Christine Blasey Ford while they were both in high school, and two further accusers later came forward (BBC, 2018b). His nomination to the Supreme Court was confirmed. The Trump Whitehouse was reported in the media as having interfered in the related FBI investigation.

<sup>7</sup> Google’s Ngram Viewer generates a graph showing the frequency of the search term in a corpus of English language books, these are likely to represent a more generalised use than the corpora of this study. (Google, 2024).

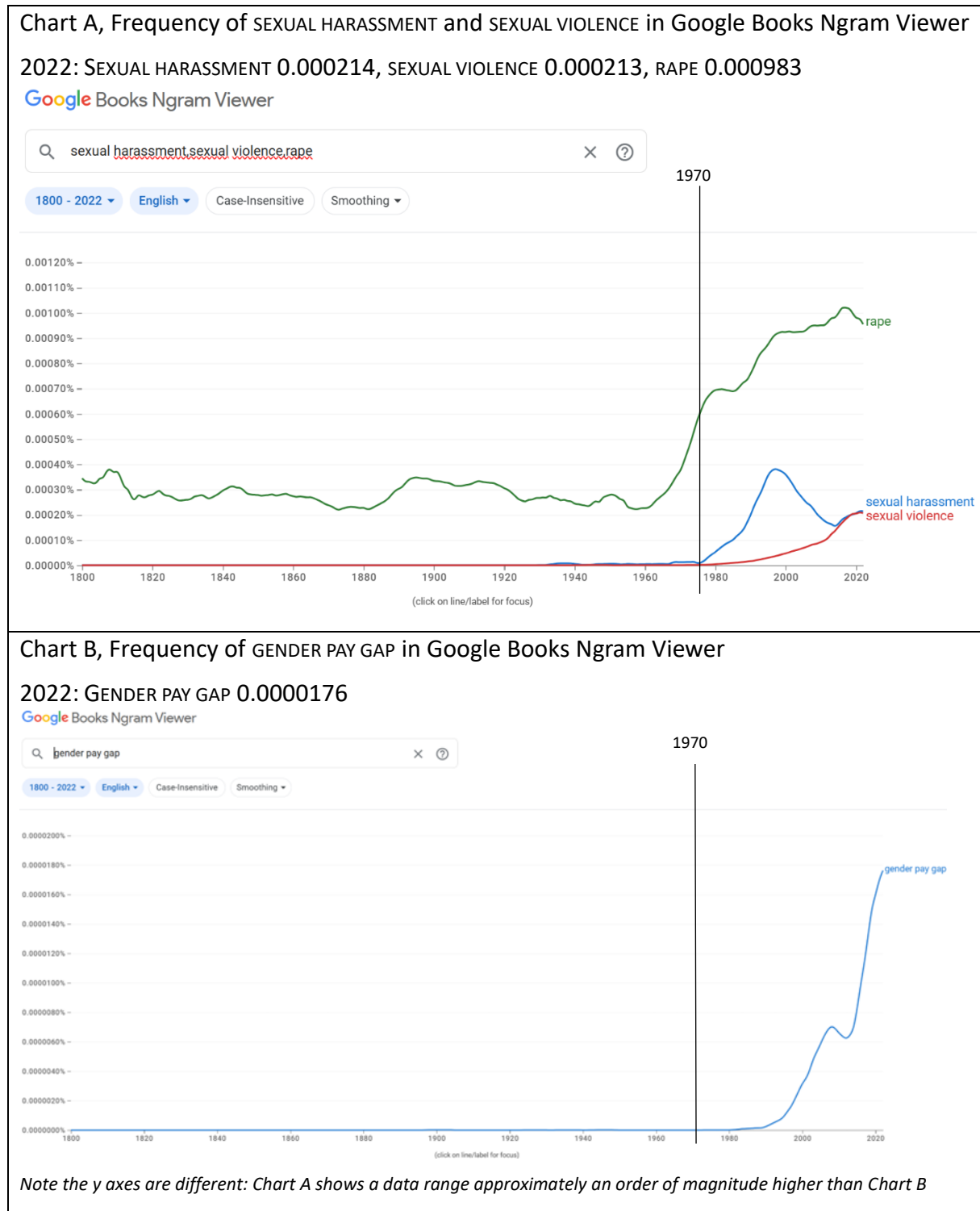
<sup>8</sup> Use of the phrase “Me Too” was started by Tarana Burke in 2006 to show solidarity to other victims of sexual abuse. It went viral in 2017 as a Twitter hashtag initially in response to accusations against Weinstein when US actress Alyssa Milano encouraged others to use it to show solidarity and highlight the scale of abuse.

# Sexual violence and the gender pay gap

## 1 Introduction

This study explores the representations within Twitter of the victims/survivors of sexual violence and those disadvantaged by the gender pay gap, using online news media text as a more traditional contrast.

Figure 1, Google Ngrams showing the changing use of terms



Targets of sexual harassment are most frequently referred to by the labels VICTIM or SURVIVOR. How individuals self-label or are labelled by others can be a sensitive issue because the word VICTIM is open to potentially pejorative meaning that could damage personal identity. VICTIM traditionally suggested that a wrong has been inflicted such as injury, mistreatment, or having been taken advantage of (Oxford English Dictionary [electronic resource], n.d.-O) however the 1989 revision adds more negative prosody “a person depreciatively regarded as helpless, passive, or failing to take responsibility”. Concurrently, the meaning of SURVIVOR shifted. Traditionally SURVIVOR referred to remaining alive through life-threatening circumstances or having the “knack of surviving afflictions unscathed”. However, the 1989 revision was broadened to include “A person who has experienced a traumatic event or past abuse, esp. of a sexual or psychological nature.” (Oxford English Dictionary [electronic resource], n.d.-bb). This change in the language now amplified by the power of social media is both influencing and influenced by an increasingly on-line society. The reciprocal relationship between language and behaviour is discussed in section 2.4.1, and its monetisation by social media in section 3.2.1.1. The potential consequences of these drivers on both issues make this topic worthy of study.

A high-level view of the study is outlined in section 1.1, and the gap it addresses in section 1.2. The research questions are introduced in section 1.3, and then section 1.4 summarises the following chapters.

## 1.1 The journey

The data for this study was sourced from Twitter and the Nexis English Language Newspapers database using the search terms SEXUAL HARASSMENT, #METOO and PAY GAP. The path through the analyses was driven by findings within the Twitter corpora. Choice of the terms VICTIM and SURVIVOR was based on their high frequencies within the corpora combined with the surprising level of dispute and emotion surrounding them. Early exploration of the pay gap corpora proved difficult as the disadvantaged are often missing from the text. Where they are referred to it is often using generic terms that do not differentiate them from others who are not disadvantaged such as *women* or *employees*.

There is a contrast between the victims and survivors of sexual violence who are generally aware of being wronged even if they have difficulty understanding or articulating it, and those wronged by the pay gap who are often unaware of their disadvantage. Within these corpora, VICTIM is more likely to be associated with negative prosody suggesting weakness, lying, and/or motivation to obtain some advantage or defame an “innocent” perpetrator. VICTIMS are blamed using traditional rape myths by others who are unconsciously protecting their view of the world as a safe place.

In contrast, SURVIVOR is more often associated with positive prosody, agency, and strength. By actively choosing SURVIVOR, individuals gain a socially preferred identity. They are more often lauded as brave, but may consequently be expected to share an uplifting version of their narrative reshaped to have a redemptive ending. The pressure to live up to the SURVIVOR label may delay personal healing, minimise the perception of harm to third parties, distance the perpetrator, and do nothing to challenge environments that normalise such victimisation.

Those disadvantaged by the pay gap are often subject to negative prosody and characterised as lacking veracity, judged as unacceptable, or viewed with contempt. Their pay gap is thus represented as a natural consequence of their choices, laziness, or arising from their labour simply being worth less. The VICTIMS, SURVIVORS, and those positioned as disadvantaged by the pay gap are the tip of the iceberg with only traces of many invisible and silenced others who might have been expected to be present.

Figure 2 evolved as a means of keeping sight of the main themes and their connections as the study progressed and is included here as an overview.

## 1.2 How this study hopes to contribute

Sexual violence is a serious crime in many countries and the gender pay gap is represented as an inequality issue, with figures for both being regularly reported at national and international levels. The reporting indicates widespread problems affecting many victims, with evidence usually suggesting that the figures represent a small minority of those actually harmed. Even in countries such as the UK and the US that have had preventative legislation for decades, progress towards eliminating these problems is very slow. This intractability suggests that there may be some mechanism(s) working to counter anti-discrimination legislation, leaving unchecked the institutional structures that enable these crimes to continue. These problems have implications for the victims beyond the immediate circumstances, with implications extending into future health and economic wellbeing. Furthermore, where such victimisation predominantly aligns with a protected characteristic, it is illegal and must be challenged. This study will use its corpora firstly to analyse representations of those who are victimised, but also to provide quantitative data to strengthen the existing qualitative theories of Feminist Legal Theory (FLT) (section 2.4.3). Additionally, this study will look for traces of those not represented directly in the text, with the aim of providing a useful insight into the part that language plays in silencing.

At the outset of this study finding existing work that was closely relevant presented a challenge. Few academic authors addressed both pay gap and sexual violence, although some journalists were raising concerns. In 2019 an unscientific ad-hoc search based on the Subject field search terms SEXUAL HARASSMENT, AND PAY GAP within the Lancaster University Library returned six documents. All of which had *work* or *industry* in the title and none were associated with linguistics or language. This suggested that the two issues may have been more often viewed in isolation and as unrelated, obfuscating structural causes and any mutually reinforcing relationship between the two.

In 2025 the same search returned 279 documents including 45 peer reviewed journal articles, all titles are shown in Table 1. Twenty-five of the journals are from the medical and health sectors. No linguistic or language titles are listed.

Journal titles	Number of articles
BMJ	8
Bmj. British Medical Journal	8
The Lancet	5
BMJ Open	2
Nature	2
Advances in Geosciences	1
Alternative Law Journal	1
American Journal of Public Health	1
Anaesthesia	1
Australian And New Zealand Journal of Public Health	1
Total	30

*Table 1, Lancaster University Library, 2025 search of peer reviewed journal  
Using search terms SEXUAL HARASSMENT AND PAY GAP listed in the Subject field.*

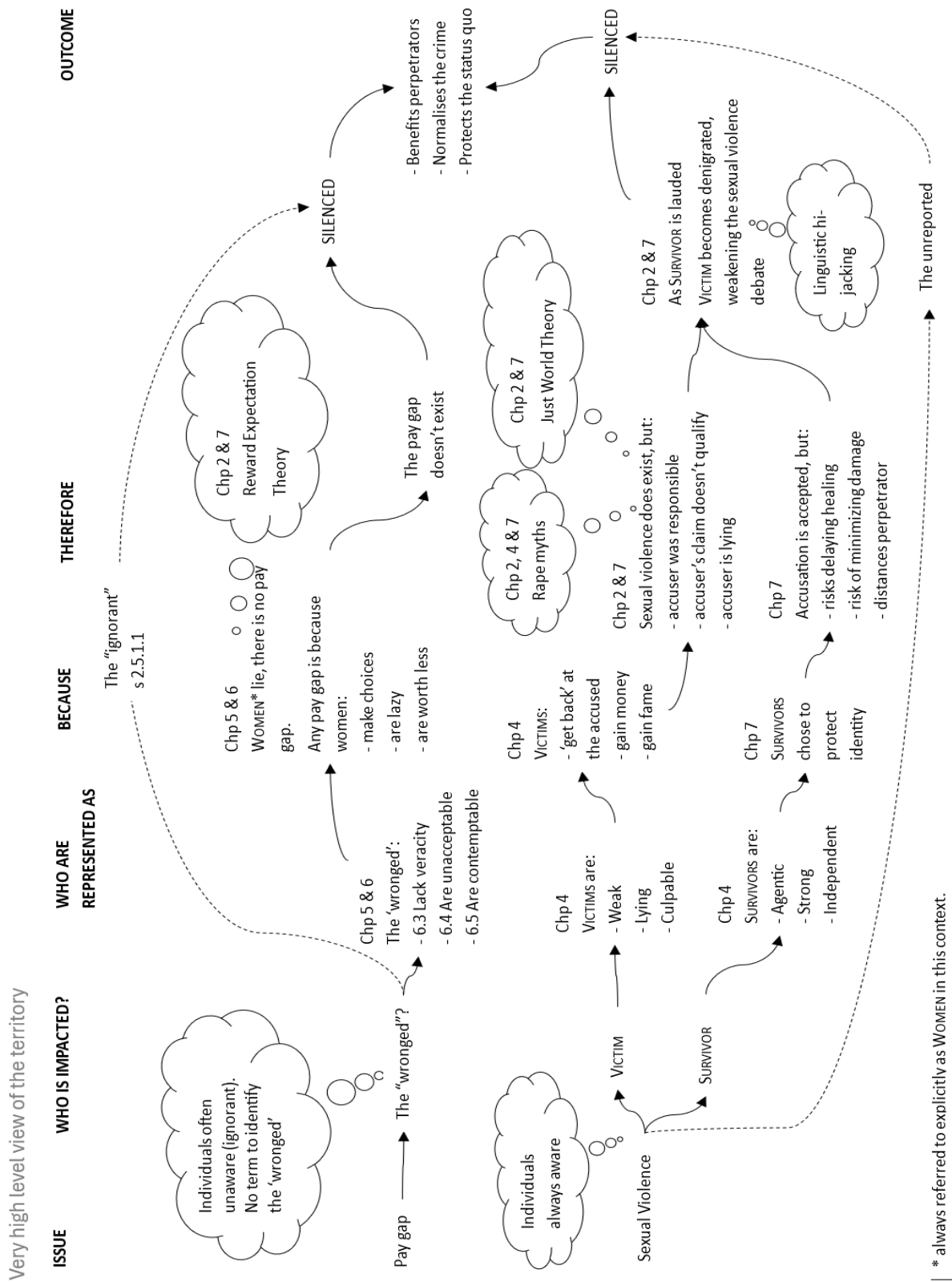
Putting aside the pay gap for a moment, a separate search using the terms VICTIM AND SURVIVOR AND CORPUS LINGUISTIC returned three documents all peer reviewed journals, but none of which were taking a corpus related approach. Messamore and Paxton’s (2021) study surveyed 3,756 organisations regarding their use of VICTIM and SURVIVOR and commented:

*“Moreover, we suggest that emergence of survivor and its linkage to victimization, gender, and especially sexual and relationship violence, remain poorly understood”* (Messamore & Paxton, 2021, p. 16)

# Sexual violence and the gender pay gap

## 1 Introduction

Figure 2, High level map of this dissertation



The increase in interest since 2019 is encouraging but suggests that the relationship between sexual harassment and the pay gap may not yet have been fully explored, particularly by corpus linguistic methods that provide quantitative data, and the present study contributes to that area. While representations of the missing cannot be counted by quantitative or qualitative means, this study will look for their traces and try to say something about them. The next section introduces the research questions.

### 1.3 The research questions

The overarching question that launched this study concerns the representation of people harmed by gendered discrimination within discourse, how this relates to the ideological landscape, and who might benefit. I address this question using Twitter corpora within SketchEngine and the manually coded datasets I take from them. I compare these with the more formal reporting language and context of news articles from Nexis. My focus is on the representations of those harmed by sexual harassment and the gender pay gap. The terms `VICTIM` and `SURVIVOR` occur at high frequencies and who is, or is not, entitled to claim either label is often contested. Hence, my first research question focuses on the negative prosody of `VICTIM` and `SURVIVOR` representations, the dominant themes<sup>9</sup> that emerge from the sexual violence concordances, and what individuals may gain by adopting a `SURVIVOR` label.

RQ1 Within the context of Twitter and Nexis sexual violence datasets, how are `VICTIM` and `SURVIVOR` represented in relation to possible positive<sup>10</sup> and negative affective responses? This question is addressed in three parts:

RQ1A, Are `VICTIM` representations more frequently negative than `SURVIVOR` representations? (sections 4.2 & 4.3)

RQ1B, What are the dominant themes associated with `VICTIM` and `SURVIVOR`? (sections 4.4 & 4.5)

RQ1C, What is gained by claiming the `SURVIVOR` label? (section 4.6 & 7)

Moving on from `VICTIM` and `SURVIVOR`, I compared the sexual violence and pay gap datasets. Starting with an analysis of the positive and negative affective responses within the pay gap dataset, I was able to compare its themes with those within the sexual violence dataset. The outcome of RQ1 analysis indicated that Twitter texts were more often negative and to a greater degree, so RQ2 focused on those, using Nexis pay gap articles as a comparison.

RQ2 How are those disadvantaged by the gender pay gap represented within tweets in relation to positive and negative affective responses, and how does this compare to representations associated with sexual harassment? This question is answered in three parts:

RQ2A Do the pay gap datasets contain more negative representations than the sexual violence datasets? (section 5.3)

---

<sup>9</sup> Themes are defined in section 3.4.2.1.

<sup>10</sup> Positive and negative are defined in section 3.4.2.

RQ2B What are the dominant themes associated with those disadvantaged by the pay gap? (section 5.4 & 5.5)

RQ2C How are the themes found in the pay gap datasets different from those of the sexual violence datasets? (sections 5.5 & 5.7)

Analysis of the RQ2 datasets revealed truthfulness and legitimacy as strong themes. RQ3 takes this forward to identify how these themes differ between the sexual violence and pay gap datasets.

RQ3, How are truthfulness and legitimacy attributed in the pay gap discourses, and how is this different from that seen in the sexual violence datasets? (Chapter 6)

Finally, I attempt to pull the strands together in RQ4 by analysing why victimised individuals might choose the SURVIVOR label and how this might relate to the wider ideological landscape. In the second part of RQ4 I try to say something about which demographics are more likely to be missing.

RQ4 How does delegitimation manifest for VICTIMS and SURVIVORS? This question is addressed in two parts.

RQ4A, Is the SURVIVOR label more beneficial than the VICTIM label for individuals, and what are the implications for society? (sections 7.3 & 7.4)

RQ4B, Who are the missing? (sections 7.5, & 7.6)

## 1.4 Summary of chapters

This section gives an overview of the material covered in each chapter.

### 1.4.1 Chapter 2: Literature review

In the literature review in section 2.2 I start by situating the public debates around sexual violence and the pay gap. In sections 2.2.1 and 2.2.2 I describe the problems of sexual violence and the pay gap, referring to reports by governmental and international bodies. I refer to the related legislation in the UK and US, and also the evidence suggesting that this has failed to eliminate either problem. Bringing together sexual violence and the pay gap, I then use reports and opinion pieces to illustrate possible links between them in section 2.2.3.

In section 2.3 I discuss my choice of tools. From the outset I took a wide-ranging corpus approach and in section 2.3.1 I describe how this was supported by the features and benefits of Corpus Linguistic (CL) and Corpus Assisted Discourse Studies (CADS). Among others I refer to Baker (2009) to show how CL is used to identify patterns of use that guide a CADS exploration of specific features. Progressing into more qualitative analysis, I discuss Fairclough's (2001, p. 21) utility in relating CDA to extratextual cultural influences. With sexual violence and the pay gap being strongly gendered issues, feminist thinking is necessarily a backdrop to this study, and I briefly describe the evolution of feminism and Feminist Post-structural Discourse Analysis (FPDA) referring to (Baxter, 2003) in section 2.3.2. I outline FPDA's view of post-feminism and neoliberalism and their influence on representations of personal responsibility and the positivity culture (Favaro & Gill, 2019; Kingfisher, 2016; Kossev, 2019; Phipps, 2019). This is followed by a description of the

political polarisation of views on sexual harassment, referring to Palomino Manjón's (2024) study on discourse on Blasey-Ford's testimony against Kavanaugh.

Section 2.4 maps out this study's broad sociological landscape, which draws from Linguistics, Philosophy, Social Sciences, Psychology, and Criminology. I start from the level of society, in section 2.4.1, by positioning language as both drawing on cultural resources and contributing to them referring to Foucault (1972), Fairclough (2001, p. 1), and Sunderland (2004, p. 8). Next, I relate language to power describing in section 2.4.2 how together they can reciprocate reinforcing the status quo to achieve specific aims (Fairclough, 2001, p. 27). In section 2.4.3 I outline Feminist Legal Theory (FLT) and its positioning of the legal system as a key part of a patriarchal state that reflects a fundamental bias against women. I then attempt to explain why individual writers might use language in ways that harm others. The theories that I draw on include: Lerner & Miller's Just World Theory (JWT) (1978, p. 1031) in section 2.4.4; Rewards Expectation Theory (RET) (Wagner & Berger, 1997, pp. 9, 12) in section 2.4.5; and Labelling Theory described in Goffman's early work (1990, p. 164) in section 2.4.6.

Section 2.5 provides some theoretical background on the lemmas *VICTIM*, and *SURVIVOR*. I start, in section 2.5.1, by introducing Nils Christie's (Sloan et al., 2018, pp. 11-22) Ideal Victim, which is used as a high standard against which the victimised are judged and how being non-ideal can lead to invisibility. In section 2.5.2 I briefly clarify that within the legal system it is necessary to use the term *VICTIM* if there is to be a perpetrator (Sinclair & Carter, 2004, p. 151). The representation of the victimised by news media and within social media is discussed in section 2.5.3, with particular reference to the Steubenville rape case and traces of neoliberalism seen within it (Barca, 2018). This section continues by discussing some of the dynamics around hashtags and the influence of social media algorithms, referring to authors such as Tranchese (2023) and Zappavigna (2017). Next, in section 2.5.4 I refer to Cole's (2021, pp. 31, 32) work describing a neoliberal backlash against the civil-rights movements of the 1960s, and the implications of this. I set this alongside Favaro & Gill's (Favaro & Gill, 2019, pp. 151-161) work suggesting that the testimonies of Jewish Holocaust survivors contributed to a shift in meaning. Referring to van Dijk and Cole (2009, pp. 14, 25), (2007, pp. 22, 138) among others, I position the label *VICTIM* as having become stigmatised. Addressing the increasing social preference for *SURVIVOR* in section 2.5.5, I describe the expectations associated with *SURVIVOR* that individuals may experience including a responsibility to heal themselves quickly, while supplying their own resources for doing so (Palmer-Mehta, 2018, p. 160). In section 2.5.6 I refer to research warning of the risks to those adopting a *heroic survivor* identity, including the potential conflict when claiming agency and requesting support (Papendick et al., 2017, p. 19 (in pdf reader)).

The practical difficulties of analysing silencing and muting are outlined in section 2.5.7, with the help of Duguid and Partington in (Marchi & Taylor, 2018, p. 39). In section 2.5.8 I describe the affect of power to regulate who may speak and who is believed, referring to (among others): Foucault's position on power, reproduced in (Medina & Wood, 2008, p. 316); and a recent high profile sexual abuse scandal (De Simone, 2024). This section concludes by explaining how a speech act can be articulated but fail to convey the intended meaning, borrowing an illustration from the activist Solnit (2019, pp. 2-4). I go on to explore how failed speech acts may originate from deep-rooted social and naturalised tropes, such as "playing hard to get" (Oxford English Dictionary [electronic resource], n.d.-w) and how this

is weaponised by the pornography industry (Maitra & McGowan, 2010, pp. 167, 168). Using a workplace study (Nielsen et al., 2024, pp. 5-7) as an illustration, in section 2.5.9 I summarise the difficulties employees may encounter in identifying a criminal act and reporting it. Staying with silencing in section 2.5.10, I introduce Anderson's (2020, pp. 2, 3) linguistic hi-jacking which operates to weaken key linguistic resources, creating what he describes as a self-correcting mechanism that maintains the status quo.

In section 2.5.11 I re-focus on the wider picture offering possible explanations for how both sexual violence and the pay gap may be framed within a right-leaning culture using inappropriately attributed agency to foreground personal responsibility, while obscuring structural causes. This draws on Fisher and Jones' (2024, p. 186) research finding that the mental health industry's focus on self-help and positivity may contribute to placing the burden of healing on the victimised.

Chapter 2 concludes by discussing the damaging impact of hate speech, in section 2.5.12, describing how it functions to silence a target and deter others from speaking out. According to Jane (2014, p. 542), and Pardo (2024, pp. 328, 329) this disproportionately affects women, and is associated hate speech with populism, neoliberalism, and "the right of the right".

### **1.4.2 Chapter 3: Methodology**

While Chapter 3 does not address any specific research question, it describes the infrastructure of the study on which the later chapters depend. My initial methodological approach, in section 3.2, describes the Twitter and Nexis data sources and justifies their selection. Using statistical sources, I go on to describe the demographics represented most often as subject to sexual violence and the pay gap.

In section 3.3 I describe the sequence of tasks undertaken to create the six corpora starting with: harvesting raw data in sections 3.3.1 and 3.3.3; deduplication and data cleaning in sections 3.3.2 and 3.3.4; aggregation of the monthly corpora in sections 3.3.6; through to the initial corpora interrogation in sections 3.3.7 and 3.3.8.

I begin section 3.4.1 with a description of van Leeuwen's Social Actor Network and Authority Legitimation Theory (2008, pp. 23, 105) and also discuss some other studies that took a similar approach, such as victims within the "refugee vs immigrant" context (Baker et al., 2008; Baker & McEnery, 2015). This section goes on to describe the manual coding of van Leeuwen's social actor and legitimation representations for 500 random concordances from each corpus. These sets of 500 concordances are referred to as the vL500 datasets. I also outline the process of assigning to the vL500 datasets two three-way flags used to identify positive and negative prosody, and veracity. In this section I also explain how these codes support my analysis and contribute to answering the research questions. The prosody and veracity codings are discussed in relation to sexual violence in section 3.4.2.1, and as relating to the pay gap in section 3.4.2.2.

From the beginning of data gathering I investigated the monthly corpora for indications of emerging themes, and I describe this in section 3.5. I then explain in section 3.5.1 the use of the VICTIM and SURVIVOR lemmas as search terms within the sexual violence corpora<sup>11</sup>. VICTIM

---

<sup>11</sup> The term *sexual violence corpora* is used to indicate all #MeToo and sexual harassment corpora.

and SURVIVOR as the subject and object of verbs were sampled, while for the pay gap corpora<sup>12</sup> the PAY GAP IS A... structure was used. The themes of silencing<sup>13</sup> and agency are discussed in sections 3.5.2 and 3.5.3 respectively.

In section 3.6 I discuss the issues encountered during this study and how these were handled. This includes how the study's prioritisation of the widest variety of language required heavy deduplication that reduced the frequency of identical and very similar occurrences, with implications for the frequencies and statistics.

### **1.4.3 Chapter 4: RQ1, Prosody and sexual violence**

Chapter 4 is based on the sexual violence datasets and aims to address research question 1. Addressing RQ1A, in section 4.2, I compare the positive/negative prosody<sup>14</sup> associated with VICTIM and SURVIVOR. Starting with some definitions in section 4.2.2, I describe how the terms *positive*, *negative*, and *neutral* are used in this study before comparing codings of the pilot batches in 4.2.3. In section 4.2.4, I describe a colleague's third party +/- prosody re-coding. Then, in section 4.2.5, I describe the results of coding the full vL500 datasets and explain the rationale for excluding concordances in section 4.2.6.

This leads into section 4.3 where I describe, in section 4.3.1, the van Leeuwen social actor pilot coding process of the vLNSHPilot<sup>15</sup> dataset, and in section 4.3.2, the third party checking of this pilot by a colleague. I explain in section 4.3.3 how I selected the van Leeuwen codes to include in subsequent analysis. Using fully coded vL500 datasets I drill down to identify the dominant themes associated with each lemma, including: Blaming/shaming, Consequences, Indifferentiation, Passivation, Speaking out, and Victim vs survivor. I explore themes associated with positive and negative prosody in sections 4.4 and 4.5 respectively. Agency is claimed to be a benefit of the SURVIVOR label and in section 4.6 I describe how this agency can become a double-edged sword. With the analysis of +/- prosody codes complete, I answer RQ1B and RQ1C in section 4.7.

### **1.4.4 Chapter 5: RQ2, Prosody and the pay gap**

Chapter 5 focuses on the pay gap corpora and research question 2. I start in 5.1 by outlining the problematic lack of any lemma that identifies those wronged by the pay gap corpora that suggests their disadvantage, and the alternative approach that was consequently required. Then in section 5.2 I describe the analysis of the pay gap datasets using examples to show how +/- semantic prosody is created. Social actor representations are then mapped to negative and positive flags in section 5.3. In sections 5.4 and 5.5 I explain using examples how negative and positive representations are created. Section 5.6 highlights rare instances of humour that I include for their rarity. I answer RQ2 in section 5.7 discussing the +/- prosody and themes of pay gap concordances and then compare them to the sexual violence datasets.

---

<sup>12</sup> The *pay gap corpora* is a blanket term for both Twitter and Nexis pay gap corpora.

<sup>13</sup> Silencing is used to obscure the truth.

<sup>14</sup> Referred to as +/- prosody in the rest of this dissertation.

<sup>15</sup> The naming scheme for corpora and datasets is provided in Appendix B.

### **1.4.5 Chapter 6: RQ3, Delegitimation**

Research question 3 is answered using the analysis of Chapter 6. I start in section 6.2 by discussing how delegitimation manifests in the pay gap discourses and how this was approached. In section 6.3 I adopt a quantitative approach to identify potentially delegitimising terms differentiating the Twitter and Nexis corpora. To provide a more granular view of delegitimation using Veracity, Acceptability, and Contempt I discuss specific examples in sections 6.4, 6.5, and 6.6 respectively. Section 6.7 compares the pay gap findings with those from the sexual violence corpora discussed in Chapter 4 to provide an answer to RQ3.

### **1.4.6 Chapter 7: RQ4, Labelling and the missing**

Chapter 7 addresses research question 4 by weighing the risks and benefits of the labels VICTIM and SURVIVOR, and also discusses those missing from the discourse.

Starting in section 7.2 I describe how the definitions of VICTIM and SURVIVOR have shifted, with VICTIM becoming denigrated and how this might correlate with a shift to the political right and neoliberalism. This leads into my discussion of the risks and benefits of VICTIM in section 7.3 explaining how this label is discredited, but facilitates access to support and justice. I explore how the word *alleged* can be problematic, and also the concept of linguistic hijacking which describes how key terms used to call out discrimination can be turned against minorities to reinforce the status quo. Turning to SURVIVOR, in section 7.4 I describe how the benefits of claiming strength and agency can bring an obligation to live up to a high ideal. While healing from a trauma involves healing over time, forced heroism may put pressure on individuals: delaying real healing; inappropriately minimising trauma; and excusing the perpetrator. Then in section 7.5 I turn to discuss those wronged by the pay gap, and for whom there is no specifically identifying label. Those missing from pay gap discourses cannot be identified directly, so I attempt a best approximation using a process of elimination to exclude the demographics who are mentioned. The individuals missing from the sexual violence corpora include those who do not report or speak out. However, the decision to speak out may be severely restricted, and this is the subject of section 7.6. Then in section 7.7 I sum up the advantages and disadvantages of the labels VICTIM and SURVIVOR and the implications for the individual, the perpetrator, and society.

### **1.4.7 Chapter 8: Conclusion**

This final chapter starts with my summary of how the research questions were addressed and the main findings of the analysis Chapters 4 to 7. Then, in section 8.2 I reflect on the study, before discussing its limitations in section 8.3, and outlining some possible areas for future studies in section 8.4. This study's contribution and implications are discussed in section 8.5 before providing some concluding remarks in section 8.6.

## 2 Literature review

### 2.1 Introduction

Both the gender pay gap and sexual violence are emotive issues that continue to be debated on Twitter and in the mainstream media, often depicting contradictory representations of the individuals involved. The VICTIM and SURVIVOR labels frequently associated with sexual violence are entirely absent from gender pay gap discourse. In fact, a key feature of the gender pay gap debate is how seldom those impacted are mentioned, leaving them silenced and missing from discourse. This presents a challenge to the researcher and necessitates a broad-based exploration of the literature.

Section 2.2 starts by situating the issues of sexual harassment and the pay gap before describing the linguistic tools employed in the methodology in section 2.3. Next, this study's sociological landscape is mapped out in section 2.4. Focus then changes to review the research into the use of VICTIM and SURVIVOR and what has been said about silencing and absence in section 2.5. Completing this review of literature, section 2.6 summarises the chapter and suggests where this study might fit in.

A number of terms are defined in this chapter:

- Sexual violence in section 2.2.1;
- Sex in section 2.2.2;
- Gender in section 2.2.2;
- Equal pay in section 2.2.2;
- Pay gap in section 2.2.2;
- Gender pay gap in section 2.2.2;
- Neoliberalism in section 2.3.2;
- Populism in section 2.3.2;
- Political right in section 2.3.2.

### 2.2 Situating the public debate

This section explains the issues of sexual violence within the context of the US and UK, the origin of most of the texts in this study, in sub-section 2.2.1, and the issues relating to the pay gap in section 2.2.2. Section 2.2.3 draws sexual violence and the pay gap together, referring to some journalistic opinion pieces pointing to possible relationships.

Sexual violence is predominantly if not exclusively a gendered issue with more reports of women suffering sexual assault than men. The gender pay gap is a statistical measure that compares women and men's average pay, so is also gendered. Hence some feminist concepts will be called upon in the following chapters and these are outlined in the following sections, but first some key definitions are given. These terms are all contested and a variety of interpretations exist, this study uses the following:

- *Neoliberalism* is used colloquially to mean a variety of things. This study uses Harvey's (2005, pp. 2, 3, 23, 65) definition of a political theory that has become a "central guiding principle of economic thought and management" which holds that deregulation of markets and trade provides the best environment for people's well-being. The neoliberal state has a minimal role beyond maintaining the structures

necessary for the market to function, and individual freedom and responsibility for personal success is foregrounded rather than facilitated or hindered by systemic factors. This world view is seen by advocates as the natural order of things.

- *Populism* is also a slippery term which Freedon (2017, p. 2) suggests may be an ideology, a mentality or a movement that overlaps with nationalism and thrives on anxiety and fear (Freedon 2017, pp. 3, 6). He explains that populists are characterised by their focus on “enemies in their midst or those who, they believe, are about to be in their midst”. While often positioned as representing the dispossessed (Freedon 2017, p. 5) populism does not necessarily promote their interests.
- *Right-wing*, as relating to politics is associated with “traditional ideas about society; relating to the belief that there should be low taxes, and that property and industry should be privately owned” (Cambridge Dictionary [electronic resource], n.d.-b). As an adjective to populism it indicates rhetoric that is “directed against an ethnically and/or nationally and/or religiously defined ‘other’. This differentiates it from the less common left-wing populism which is not ethnically exclusive (Pelinka, 2013, p. 7).

### 2.2.1 Sexual harassment/violence

The present study uses the term *sexual violence* when discussing the spectrum of unwanted sexual acts which is defined by the World Health Organization (WHO) as:

*“Any sexual act, attempt to obtain a sexual act, unwanted sexual comments or advances, or acts to traffic or otherwise directed against a person’s sexuality using coercion, by any person regardless of their relationship to the victim, in any setting, including but not limited to home and work.”*

*(Garcia-Moreno et al., 2012, p. 2)*

Within the TwitterMT\_1Yr corpus *SEXUAL* modifies *harassment* (638, LogDice 12.5) more frequently than *violence* (160, LogDice 10.7), this reflects that incidents are more commonly described using verb forms like *harassed* and *harassing* than using *subject to sexual violence*. Consequently, this study uses the term *SEXUAL HARASSMENT*. However, *sexual violence* is also used as the over-arching term as defined by the WHO.

While sexual violence can be perpetrated by and upon anyone, it is a gendered crime with male defendants making up 98% of prosecutions for sexual crimes in the UK in the year ending Dec 2017 (Ministry of Justice, 2018, Table 2).

In a BBC interview (Gill & Rahman-Jones, 2020) Tarana Burke, founder of the MeToo movement, described the systemic and pervasive nature of sexual violence and its impact on people of all demographics, particularly “All the people who are not rich, white and famous, who deal with sexual violence on an everyday basis. We need to talk about the systems that are still in place that allow that to happen”. Sexual crimes often take place in private with no witnesses and little material evidence so are liable to disintegrate into “he said, she said” cases. This deters many victims from reporting crimes, and for those who do report the conviction rates are poor. Where a case rests on whose testimony is upheld as truth, the erosive delegitimation caused by rape myths is particularly damaging (Crown Prosecution Service, 2021). Hersch takes down a particular myth saying that sexual harassment is not

associated with desire but is used to intimidate and exercise power over others (2011, p. 630).

The BroadBand Commission for Digital Development Report (BCDD) (2015, pp. 2, 22) addressed the concern that online communication technologies and social media presented new opportunities to harm women and girls. They attribute gendered cyber violence<sup>16</sup> to societal barriers and also to perpetrators being seldom held to account and consequently feeling uncountable. The BCDD report (2015, p. 15) describes the online abuse of women and girls (not exclusively sexual):

- “women are 27 times more likely to be abused online than men”;
- “61% of online harassers were male”;
- “73% of women have been abused online”;
- “Women aged 18 to 24 are at a heightened risk of being exposed to every kind of cyber VAGW; they are “uniquely likely to experience stalking and sexual harassment...”.

At the end of the data gathering phase for this study the “Prevalence and reporting of sexual harassment in UK public spaces report” by the All Party Parliamentary Group (APPG) (Advance Pro Bono, 2021) which was based on a YouGov survey, found that the situation had not improved, and:

- “71% of women of all ages in the UK have experienced some form of sexual harassment in a public space.” (p. 6);
- “3% of 18-24 year-olds specified that they had never experienced any of the listed forms of sexual harassment” (p. 11);
- “Over 95% of all women did not report their experiences of sexual harassment” (p. 17).

The APPG report also found an attitude that women were expected to accept an unspecified level of sexual harassment (Advance Pro Bono, 2021, p. 19).

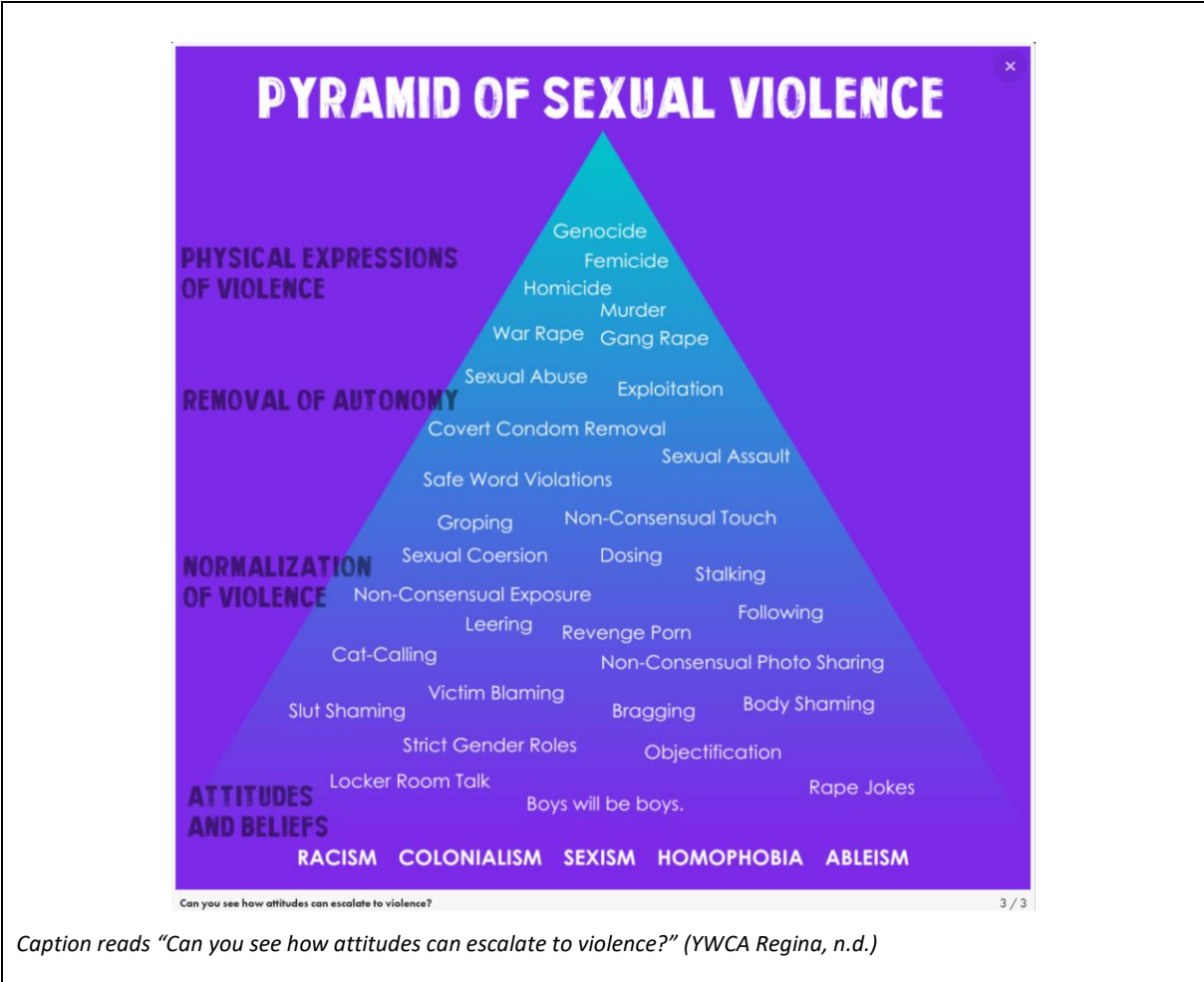
Within Twitter the calling-out of “less serious” non-physical offences is debated, and specifically whether it harms “more serious” cases such as rape, or those involving children. This study takes the view that all degrees of sexual violence are unacceptable. No one would argue that stealing was acceptable provided losses were low value, and where on Figure 3 would the threshold of acceptability be drawn?

Within the data of this study the terms *sexual harassment*, *assault*, *consent*, and even *rape* appear to be frequently misunderstood and misused. This may contribute to real subjective differences of perspective, for example at the less serious end of the scale, on whether a particular instance was clumsy flirting or sexual harassment, a friendly hug or an unwanted grope.

---

<sup>16</sup> Cyber violence is defined as “dehumanizing, aggressive and harmful acts that are in turn physical, psychological, sexual, and exploitative”. Moreover, “Violence online and offline, or ‘physical’ VAWG and ‘cyber’ VAWG, feed into each other.” and often reinforces existing power relations. (The BroadBand Commission for Digital Development, 2015, pp. 5, 7, 26).

Figure 3, The sexual violence escalation pyramid



**2.2.2 The pay gap**

The terms used in pay gap discussions are often confused, however this study uses the following definitions<sup>17</sup>:

- *Gender* follows the UK Census 2021 guidance published in 2025: "Gender identity refers to a person's sense of their own gender, whether male, female, or another category, such as non-binary. This may or may not be the same as their sex registered at birth." (Office for National Statistics, 2025a, Glossary). The definition of gender has been contested in law and elsewhere.
- The UK Census 2021 defines *sex* according to how a person was registered at birth, either "Female" or "Male". (Office for National Statistics, 2025a, Glossary)

<sup>17</sup> The pay gap is often referred to as the Uncontrolled pay gap in the US, while the expression Controlled pay gap removes variables such as gendered career segregation and glass ceilings to come close to equal pay.

- *Equal pay* (or wage gap) is defined as individuals receiving equal remuneration for work that is of equal or similar value, or makes similar demands (*The Equality Act 2010 (Gender Pay Gap Information) Regulations 2017*, 2017, part 5, chapter 3, section 64, 65).
- The *pay gap* is a statistical calculation of the “difference between average hourly earnings (excluding overtime) of men” (Office for National Statistics, 2023). It is applied to a whole population, an industry, or company, but says nothing about the situation of an individual. Consequently, the pay gap includes equal pay, but also wider issues with socially constructed barriers (Government Equalities Office, 2019). Pay gaps exist between many demographics.
- The term *gender pay gap* is used colloquially and in official reporting to refer to women and men. While it could be phrased more accurately as the sex pay gap, this could be misinterpreted as obtaining pay for sex.

In the US equal pay was mandated into law in 1964 (U.S. Equal Employment Opportunity Commission) and an end to violence against women followed in 1994. This was paralleled in the UK by The Equal Pay Act (1970) and The Equality Act (2010). Despite some progress this legislation has not eliminated failures of equal pay, the pay gap or sexual violence.

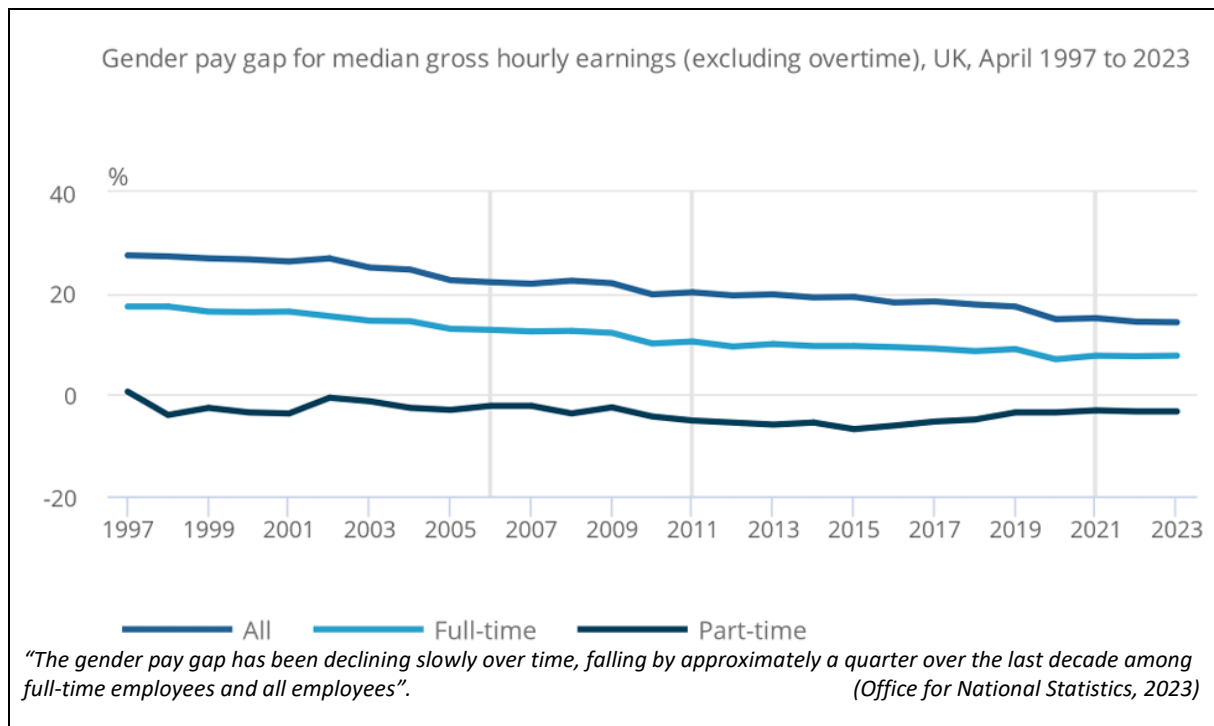
Pay gaps exist for a number of demographics including for example: ethnic, disability, and LGBTQ+ groups. Researching any of these would be worthwhile, but this study focuses on the gender pay gap which is the most frequent in these corpora.

The UK Office for National Statistics (2023) reports that the pay gap is closing, but at a glacial and slowing pace, see Figure 4. Furthermore, after controlling for improvements in women’s educational achievements, pay equality in the UK has barely progressed in the last 30 years (Andrew et al., 2024, p. 297). The ONS found the part-time pay gap was slightly positive with women earning more than part time men, however the issue here is the disproportionate number of women holding part-time roles, which command lower hourly rates than full time working (Office for National Statistics, 2023, pp. 3, 4).

## Sexual violence and the gender pay gap

### 2 Literature review

Figure 4, ONS gender pay gap 1997 - 2023



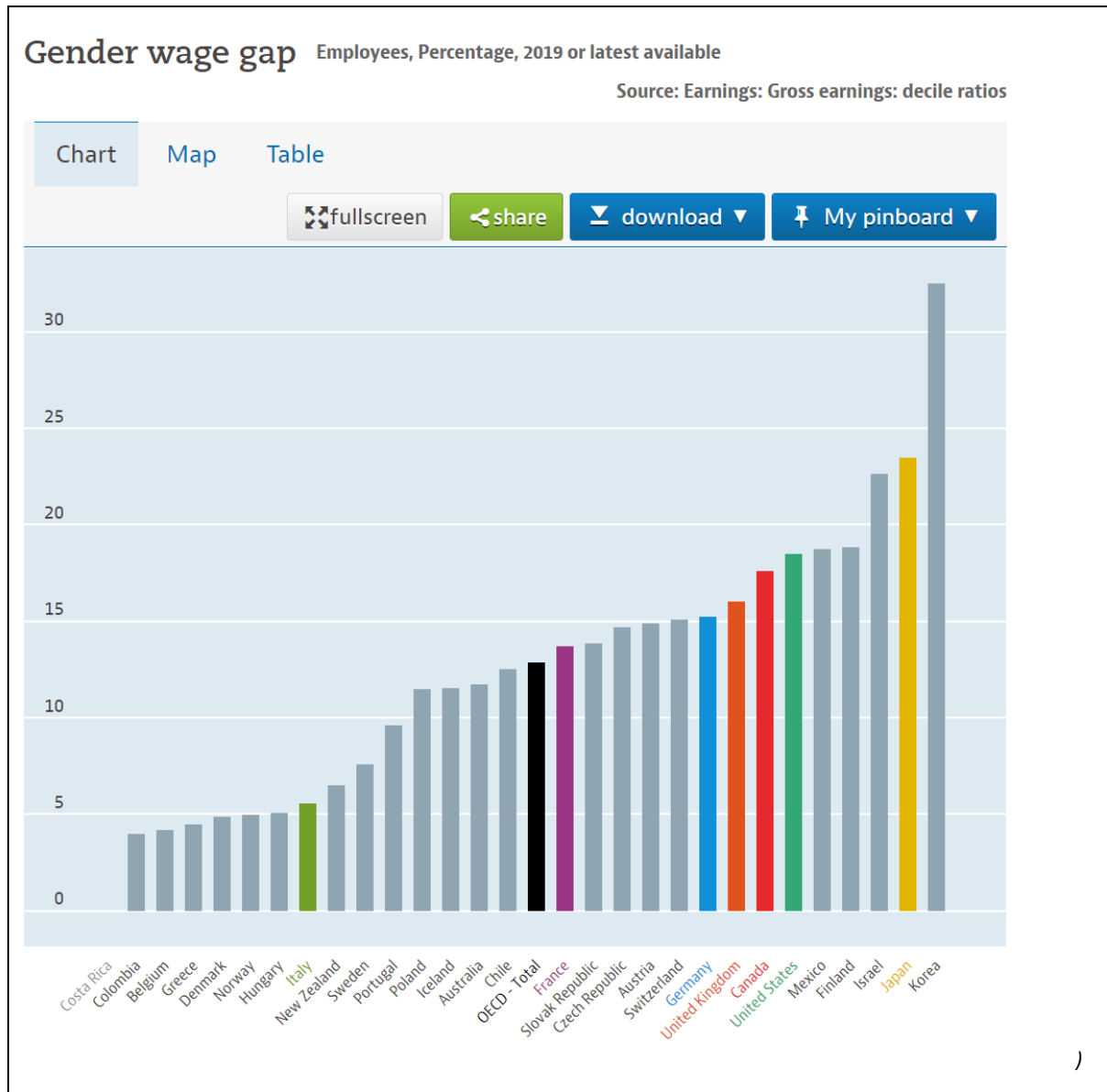
The UK government's 2019 "Gender Equality Roadmap" (Government Equalities Office, 2019, pp. 6-8) reported their findings on the drivers of gender inequality. They identified the following causes, none of which suggest that solutions are under the control of the individual:

- limiting attitudes towards gender from a young age; (p. 6)
- career segregation into lower paid occupations; (p. 7)
- failure of the working age benefit system; (p. 7)
- time out of work caring for children; (p. 8)
- other unpaid work for others; and (p. 8)
- barriers to returning to work. (p. 8)

Increasingly countries have introduced mandatory gender gap reporting such as the UK in 2017, enabling employees to more easily identify and challenge inequalities.

Figure 5, OECD gender wage gap in 2019

(OECD, 2020)



After more than five decades of legislation, failures of equal pay continue to emerge, for example: Birmingham Council (2012-2024) (Headley & Rack, 2024); and Asda (2014-2025) (Butler, 2025). Some employers forbid discussion of wages, while at the same time some employees may view disclosing salary information as inappropriate (Dominus, 2020). Failures of equal pay remain invisible when underpaid and un-unionised workers are unable to compare salaries and challenge injustice. According to The Guardian newspaper, Norway improved its pay transparency by the annual publication of its citizens' tax records, and this is often represented as driving Norway's smaller (5<sup>th</sup> lowest) gendered pay gap (Collinson, 2016), see Figure 5. However, the mechanism may be more complex than that as Finland, which also publishes tax returns, has the 4<sup>th</sup> highest gender pay gap. Barry (2018) warns that taxable income may not provide total transparency as it is not the same as salary.

Equal pay figures are improved by a legislated minimum wage which ensures that women and men are paid the same in the lowest paying jobs. However, industries where high numbers of low paid roles are predominantly occupied by women will have a larger gender pay gap (Office for National Statistics, 2019). This is often related to the *motherhood penalty*, where women need part time or flexible work to accommodate unpaid care work. However, women do not deliberately choose roles below their skills, experience and previous pay grade (Bennett, 2024, p. 347). In fact the multiple causes of the pay gap are largely beyond deliberate wage discrimination, involving career segregation, recruitment and promotion, and may also be associated with sexual violence in the workplace (Shaw et al., 2018).

Nardone (2018, pp. 179, 181) conducted a cross-linguistic German/Italian study providing evidence of both difficulties obtaining equal pay, and barriers to gaining employment. Difficulties arising from unpaid care work featured in both languages' corpora. She found evidence that the culture with the higher rate of female employment (Germany) returned higher frequencies of pay gap related modifiers, than the culture with a lower rate of female employment (Italy) where there were higher frequencies of modifiers related to barriers to obtaining work (2018, p. 183). Nardone observed that the problems in society were appearing in the language.

Equal pay in sport featured prominently, particularly around the 2019 world cup which fell during the data harvesting period. Analysing equal pay tweets from a single month after the Women's World Cup competition, Alkhamash (2021, p. 31) found that those with positive sentiment were associated with pride and direct support for the US women's team, while those negative sentiment related to challenging the US Soccer Federation (USSF) seen as responsible for pay disparity. Hence both positive and negative sentiments represented the unequal pay as a bad thing, and none challenged its existence or validity. She identifies barriers faced by women in sport including less media coverage, and bias in how women are named, partly attributed to under-representation of women in newsrooms.

Even when individuals can identify injustice redress for claims of un-equal pay and harassment need to be pursued by the individuals themselves through internal company grievance procedures, a process that can be traumatic, expensive, and not necessarily successful (Cooke, 2024).

### **2.2.3 Links between sexual violence and the pay gap**

In her commentary for the Institute of Fiscal Studies (IFS) Deaton Review Cooke suggests that the level of women's economic and political power is "precursor and consequent" to reducing domestic violence and sexual harassment (2024, p. 354). Similar opinions have previously been expressed by politicians and journalists. In her International Women's Day speech to the UK Parliament the Shadow Minister for Women and Equalities, Dawn Butler (2020) referred to the "intractable structural inequalities" and "continuing violence against women and girls" over the fifty years since preventative legislation was enacted in the UK. Ammerman and Groyberg in *The Harvard Business Review* (2017, pp. 1, 2 (in pdf reader)) went further, suggesting that women in all types and levels of workplace can be subject to sexual harassment that operates to keep women "in their place", frustrating career paths and reducing the number of women in senior management who would have the power to change workplace culture. In the magazine *Marie Claire*, Krawcheck (2020, pp. 2, 3 (in pdf

reader)) described female Wall Street employees “trapped in an abusive work environment” where “sexist treatment has driven numerous others out of their jobs, negatively impacting their earning power and their wealth”. She went on to decry reporting routes for grievances through managers predominantly from a demographic unlikely to have suffered either sexual harassment or career barriers personally, and whose workplace experience normalised such behaviour. More recently, in an opinion piece on *The Conversation*, Singh referred to the poor retention rate of women in IT as being due to “minor to severe harassment, sexism, discrimination and misogyny to explicit death threats. Their expertise is challenged, their contributions are not well-received and their roles are diminished.” (2023). Fuhrmans, in another *Wall Street Journal* opinion piece quoted management experts and executives as saying that “...harassment can be a direct side effect of a workplace that slights women on everything from pay to promotions, especially when the perception is that men run the show and women can’t speak up” (2018, p. 1 (in pdf reader)). The same year on CNN Carpenter put it more bluntly “Harassment is about power dynamics - and in many jobs, money equals power.” (2018, p. 1 (in pdf reader)).

Throughout this chapter and the whole study there is a thread of tension between personal responsibility and issues inherent within structural frameworks. Contested situations become visible through discourses advising women to take action to avoid sexual violence, or to improve their negotiating skills to eliminate the pay gap. While these personal actions may or may not protect an individual they do not resolve the structural problem, suggesting that personal actions are not the cause of the injustice, echoing the UK Gender Equality Roadmap (Government Equalities Office, 2019) discussed in section 2.2.2.

### **2.3 Methodological linguistic tools**

Having situated the issues, this section describes the main tools used in this study: corpus linguistics (CL) chosen to deal with the large volume of electronic texts (section 2.3.1); Corpus Assisted Discourse Studies (CADS) to allow access to resources beyond these very brief and often incomplete messages (section 2.3.1); and Feminist Post-structural Discourse Analysis (FPDA) which embraces multiple and contradictory voices (section 2.3.2).

#### **2.3.1 CL and CADS**

Baker (2010, pp. 10-11) describes corpus linguistics (CL) as having a number of benefits including quantification of the relationships between words and enabling unexpected features to arise from corpora. This was a lesson from the early days of CL when findings were seen to differ from researchers’ intuitions (Sinclair, 1991, p. 4). CL’s computerised statistical methods also enable quantitative comparison of different corpora, for example texts from different sources, times, or topics (Baker & McEnery, 2015, p. 5). This is a useful feature for this study which has six corpora.

Approaches to corpus research are often discussed as binary choices between either corpus-driven or corpus-assisted approaches (McEnery et al., 2006, pp. 8-11). A corpus-assisted researcher brings intuition and expectations to the corpus and searches the data for evidence to support their hypothesis. In contrast, a researcher pursuing a corpus-driven approach starts with no preconceived idea of an outcome, first generating statistical information to identify the features to be examined (Baker, 2010, p. 109). Thus a corpus-driven approach avoids the accusation of introducing bias (Baker, 2014, p. 17). However,

some bias is inescapable as the researcher is likely to choose a topic that they have an interest in and therefore about which they are not totally naïve (McEnery et al., 2006). For the present study CL is particularly interesting for its ability to identify the absence of features that there was good reason to expect (Duguid & Partington, 2018, pp. 39, 42). Such expectations could arise from previous research or from the researcher's knowledge of the social environment gained in a previous more qualitative analysis (Duguid & Partington, 2018, p. 43).

The large corpora of CL bring issues of duplication. This can arise from either boilerplate text applied to a news article, or parts of the story being repeated across different editions, building duplication into the raw data (Baker, 2014, p. 160). Within Twitter data the large numbers of very short tweets include high percentages of retweets and copying, which also drives duplication. Regarding the similarly short texts of small ads, Baker concluded that even brief messages, shorter than many tweets used in this study, were sufficient to identify linguistic patterns (2014, p. 162).

Critical Discourse Analysis (CDA) is a qualitative tool that enables sociological drivers to be linked to features in the language. Fairclough (2001, p. 21) suggests that the search to discover such extratextual influences on language motivates CDA researchers to "develop self-consciousness" about the roots of discourse that extend deep into ideological assumptions (2001, p. 139). Using CDA alongside a corpus method to study language for those "discourses that permeate the relationships among social groups, particularly those involving imbalanced power dynamics" brings synergistic benefits according to McEnery et al (2015, p. 238). Thus CL and CDA can be combined in an iterative process harnessing the strengths of each, allowing the patterns identified by CL to be analysed using CDA methods to reveal what is "implied, inferred, insinuated or latently hinted at" (Baker et al., 2008, pp. 295, 296). Conversely, unexpected features in concordances can be investigated using CL to establish statistical relevance. Transitions between statistical exploration and qualitative CDA analysis can then be informed by reasoned, if still subjective, decisions (Baker & Levon, 2015, p. 222).

The present study uses Corpus Assisted Discourse Studies (CADS) which is designed specifically to answer social questions, and when in combination with a critical discourse element, is well suited to study inequalities and power (Gillings et al., 2023, pp. 4, 5). Gillings et al describe CADS as drawing on a range of corpus linguistic methods that are used alongside CDA. CADS studies then, can exist on a spectrum from completely corpus oriented to completely discourse analysis focused, with the strength of the resulting synergy dispelling the notion of an either/or dichotomy of methods.

Real world data can be messy and CADS allows researchers to adopt a responsible "craft attitude" (Gillings et al., 2023, p. 55) by combining tools flexibly and viewing research as a process rather than neat products. Gillings et al advise researchers to: embrace uncertainty; allow for recursive design; and see research as a narrative process (2023, p. 55). Alongside this flexibility they warn against overconfidence in CL and for maintaining a self-reflexive attitude to research to ensure "evidence is neither over- nor under-interpreted" (Gillings et al., 2023, p. 54).

### 2.3.2 The feminist context and FPDA

Sexual violence is predominantly, if not exclusively, a gendered issue with four and a half times more reports of women suffering sexual assault than men (Office for National Statistics, 2025b). The gender pay gap is a statistical measure that compares women's pay to men's so is also gendered (Office for National Statistics, 2023, pp. 3, 4). Some feminist concepts will be called upon in the following chapters and these are now outlined in this section, starting with a little background.

Across the 1960s and 70s, feminism focused on the power of language as central in bringing about discrimination and as a means of fighting it (Kramer, 2016, p. 65). Initial focus was on the more visible and obvious features such as: the higher frequency of male pronouns in texts; the convention of putting male before female pronouns; and the generalised use of male pronouns for humankind. These features were challenged, resulting in reduced use of gender biased language. Concurrently, the expression *sexual harassment* became common currency to identify behaviours, now viewed as wrong, that were previously normalised as flattering or even funny (Kramer, 2016, p. 65). Overtly sexist language was challenged resulting in it becoming less common, however it may have been replaced with more subtle versions, for example using modifiers such as *alleged* and *real* (Tranchese, 2023, pp. 164, 108, 109), which are discussed in Chapters 4 and 6. The mid-20<sup>th</sup> century feminists bringing about these changes identified universal causes as underpinning women's issues, however later decades saw this view challenged. Increased awareness of intersectional demographics such as ethnic and LGBTQ+ groups, pushed theorists towards a more nuanced model in both the context of sexual violence (Advance Pro Bono, 2021) and the pay gap (Gould et al., 2016, pp. 10-13).

From the 1990s, with equal pay legislation secured within several countries, post-feminists declared that equality had been achieved, feminism was redundant and irrelevant, and women were now free to pursue personal ambitions (Lawson, 2023, p. 10). Thus post-feminism connected liberal values with women's choices and autonomy (Banet-Weiser, 2019, pp. 171, 172). Unconvinced, non-post-feminism feminists continued to maintain that there was still work to be done.

As social media platforms gained momentum the narrative was taken up by "very visible, mainly white, public individuals" creating a new form of feminism (Banet-Weiser, 2019, p. 173). This brand of feminism focused on individual celebrity crimes like those against Weinstein<sup>5, page 1</sup> and R Kelly<sup>18</sup> rather than challenging the culture that makes such crimes possible. A characteristic of this individualisation is the "insidious flip" of a "deny-and-then-accuse" pattern that counter-accuses the original claimant(s) of defamation (Banet-Weiser, 2019, pp. 174, 175). Despite its new more celebrity-driven focus, and instant access to huge audiences, this new brand of feminism has failed to deliver tangible changes for women's rights. Conversely, in the US the right to safe and legal abortion has been lost, Title IX amendments increase the barriers college students face when calling-out sexual violence, affirmative action for diversity ended, and online there has been a backlash against #MeToo and women's rights more generally (Alison Phipps, 2019, p. 65). In the 28 pages of the

---

<sup>18</sup> RK refers to R Kelly (a US music celebrity) convicted of sexual violence crimes in 2021.

“Women’s Rights in Review 30 Years After Beijing” report (UN-Women (United Nations Entity for Gender Equality and the Empowerment of Women), 2025) the word *backlash* appears 10 times.

Feminist Post-structural Discourse Analysis (FPDA) formed as a reaction to the monolithic and universal causes of mid-20<sup>th</sup> century feminism that were seen as affecting all women. It is a methodology designed for face-to-face, turn-taking, and dynamic speech events (Baxter, 2003, p. 65). Hence, its methods are not a perfect match for this study but some of its explanations, particularly its linking of post-feminism to neoliberalism will be called on. While similar to CDA, a key difference is FPDA’s view of power as constantly shifting between interlocuters who may experience different degrees of power in different areas and times of their lives (Baxter, 2003, p. 73). FPDA visualises discrimination as articulated by multiple marginalised voices (Baxter, 2003, p. 72). Thus Baxter urges FPDA researchers to embrace multiple and contradictory voices, acknowledge the constructed nature of all aspects of research, and warns against seeking definitive closure when analysing meaning (2003, pp. 45, 61). Within the present context where social media is ubiquitous, Bouvier highlights that the wide variety of meanings associated with a single hashtag may be “diverse, complex and sometimes indistinct”, and connected only by the effect of buzzwords, symbolism, and outrage (Bouvier, 2022, pp. 192, 193). In contrast to older technologies however, social media may be showing us the wider range of ideological views that co-exist from the non-elite parts of society (Bouvier, 2022, p. 180). From the end of the 20<sup>th</sup> century a neoliberal trend was identified within developed nations, characterised by the repositioning of social welfare as: firstly no longer affordable; and secondly, disempowering for recipients (Kingfisher, 2016, p. 258). Situated in a period of widespread economic recessions and austerity, this led to down-scaling of welfare programs at a time when there was greatest need (Alison Phipps, 2019, p. 63). Thus, right-leaning regimes prioritised personal responsibility with individual solutions undermining potential for collective action (Kingfisher, 2016, pp. 259, 260). As welfare programs shrank NGO support services filled the vacuum but their funding models exposed them to influence from government and market pressures (Lamphere, 2016, p. 55). Consequently, support services have also promoted personal responsibility and self-help, representing clients as free to make unconstrained autonomous choices. Idealised as “active, entrepreneurial agents” neoliberal citizens are expected to re-make themselves following personal “failures” (Kingfisher, 2016, pp. 260, 261). Favaro and Gill suggest that women’s generalised re-making of self is reinforced by a consumer culture of self-improvement advertisements and editorial content by women’s magazines (Favaro & Gill, 2019, pp. 157, 158).

The activity of “active, entrepreneurial agents” remaking themselves, which maps more closely to the definition of SURVIVOR than VICTIM, could be reframed as neoliberal “compulsory hopefulness and optimism” (Favaro & Gill, 2019, p. 154). This may carry a burden of expectation of the victimised healing themselves quickly, and returning society to its previous productive and economic position (Favaro & Gill, 2019, p. 155; Kosew, 2019, p. 182). Within Twitter, support for speaking out on sexual violence has been associated with the political left, particularly Democrats in the US for example in relation to Dr Blasey-Ford’s testimony against Brett Kavanaugh (Palomino Manjón, 2024, p. 106). In response, Republican politicians positioned the #MeToo movement as a political tool being used against men and to harm the Republican party. They used Twitter to discount Blasey-Ford’s claim as bogus,

positioning Kavanaugh as the real victim (Palomino Manjón, 2024, p. 119). In turn, Republican party members were constructed as misogynists (Palomino Manjón, 2024, p. 115). Perhaps surprisingly, some female supporters of Kavanaugh distanced themselves from and demonised the accusers, #MeToo, and the feminist movement (Palomino Manjón, 2024, p. 119). A possible explanation was offered by Criado-Perez (BBC, 2014). Upon winning a court case against her trolls, one being a woman, she said “I don't see why we should think that women who are brought up in a society steeped in misogyny should be any less affected by it and any less likely to hate women”. Palomino Manjón (2024, p. 106) also observed the variety of positions taken, noting that both her hashtags #KC and #NoKC were used by those supporting both Kavanaugh's appointment and Blasey-Ford's testimony that challenged his appointment.

The next section attends to the part that language plays, and then why individuals might be motivated to doubt and judge accusers negatively. Literature regarding both the gendered use of language around VICTIM and SURVIVOR, and also the gendered nature of the lived experiences of both sexual violence and pay disparity, is discussed in section 2.5.

## **2.4 The sociological landscape**

CADS studies often draw on theories from other disciplines relevant to the topic (Gillings et al., 2023, p. 5). Studies underpinning our understanding of the complex relationship between language and the social environment are discussed, and here these are drawn from the Social Sciences, Psychology, and Criminology. Sections 2.4.1 and 2.4.2 outline the early academic work on how language interacts with culture. Feminist Legal Theory and its conception of the powerful and gendered role of the legal system is introduced in section 2.4.3. Possible mechanisms for why and how this might operate within the social world are then discussed in section 2.4.4 in terms of how people internalise the Just World Theory (JWT) to feel safer, and in section 2.4.5 by proposing how Rewards Expectation Theory (RET) may contribute to the gender pay gap. Section 2.4.6 discusses Labelling Theory's impact on individuals.

### **2.4.1 Discourse is reality, even when it's not real**

As early as 1902 Cooley (cited in Fredrickson & Roberts, 1997) suggested that an individual's self-image is socially constructed and underpinned by their interpretation of how others see them, which he termed the looking-glass self. During the latter half of the 20<sup>th</sup> century post structuralist philosophers proposed that rather than concrete facts and reality that were fixed for all time, we shape our own reality through the discourse that we use (Foucault, 1972, p. 54). This has the consequence that harm can be inflicted by damaging discourses. These ideas are important because of two serious implications. Firstly, discourses are “practices that systematically form the objects of which they speak” (Foucault, 1972, p. 54). Secondly and more specifically, those objects include peoples' “looking-glass” selves: for example, if I think that you hate me I can feel wounded. Taking this a step further, Sunderland (2004, p. 8) suggested that this could happen through “ideological ‘flows of information’ that construct the world through language and texts, and subject position individuals”. This implies that information being communicated may too readily be accepted as being correct or normal, when perhaps it should be open to question. The circulation of ideological flows allow the continuation of their damage, while interlocutors remain

unaware. Fairclough (2001, p. 33) suggests that people may remain ignorant of the legitimating and delegitimizing effects of “particular power relations” and their social importance. Thus, information that is conveyed implicitly through tweets could legitimise or delegitimise, but without either writer or audience being aware.

### **2.4.2 Language and cultural change**

Companies spending money on advertising clearly believe that well-chosen language drives real world changes in behaviour. A recent lab-based study into the effects of marketing language (Fritze et al., 2024, p. 23) reinforces the belief that changes in branding language preceded changes in purchasing behaviour, rather than the other way around. Much earlier, Butler suggested that language change may precede societal change (1999, pp. 120, 124), for example being labelled a *girl* from birth causes a child to become “girlled over time”. While this makes language a powerful tool for the people using it, the degree of power involved is related to the speaker/writer’s authority and circumstances (Baker et al., 2008, pp. 280, 281). This suggests that how we address individuals, or they refer to themselves, is important. However, it is not as simple as just talking or writing a change into being.

The 1974 book, *Language and woman’s place*, (Lakoff & Bucholtz, 2004, p. 52) concerns itself with various aspects of gendered asymmetry in language. Lakoff and Bucholtz highlighted earlier attempts by some women’s magazines to improve the status of women by promoting new labels of “homemaker”, “household executive”, or “household engineer”, in preference to “housewife”. They noted that none of those terms were adopted as general use, concluding that for language to drive social change, society must be receptive, and even then the process will be slow and indirect (Lakoff & Bucholtz, 2004, p. 72). Fairclough agrees that language may change the real world, but cautions that the likelihood of change is related to the degree of power that the speaker has, and the beliefs, values, and experiences circulating within the culture at the time (Fairclough, 2009, pp. 518, 520).

The legal system is a major arbiter of power, and the next section discusses its operation from a feminist perspective.

### **2.4.3 Feminist Legal Theory**

Feminist Legal Theory (FLT) is a broad field of research that challenges the view that the legal system is neutral. Instead it claims that the system has built-in inequalities because it was created by and operates within a patriarchal society. Thus the legal system is orientated to a male perspective that may contradict women’s realities, and hinder their attempts to be heard and obtain justice.

Debate in the late 1980s was dominated by the publication of *Toward a Feminist Theory of the State* (MacKinnon, 1989) and *Feminism and the Power of Law* (Smart, 1989). MacKinnon and Smart agree on many of the challenges women encounter in accessing justice and equality, but differ on what they see as the causes and therefore what solutions could be effective.

Smart (1989, pp. 6-25) foregrounds Foucault’s thinking on discourse, power, and truth and sees the legal system as based on patriarchal rationalisations and knowledge that have resulted in construction of gendered laws that are presented as neutral and often accepted as being so. However, from Smart’s (1989, p. 26) perspective sexual violence is narrowly

defined in law and shaped by male-centric legal system discourses, and this definition disqualifies other definitions that might capture women's experience. Smart (1989, p 5) views attempts to remove bias against women solely by implementing further laws as trying to change the system using the system's own biased affordances. Thus, Smart (1989, p. 25, 165) recommends that in addition to changes to laws, feminist activists should de-centre the law by challenging its power more broadly, and also pursue other types of feminist activism.

In contrast, MacKinnon (1989, p. 4) positions feminism as analogous to Marxism, in that she envisages the institutions of the state as working together to maintain the legal system as an expression of male dominance. This gives rise to structural barriers that disadvantage some groups, and MacKinnon (1989, p. 5) criticises liberal values that represent these inequalities as personal issues. Sexual violence is central to her theory, operating to enforce gendered power relations (MacKinnon, 1989, p. 61). She describes conflicting gendered perspectives on what constitutes sex, and sexual violence, with the law interpreting individual accounts of sexual violence solely from the male perspective. MacKinnon believes that improved justice for women can and must be delivered by changes to the structure of the legal system.

In regard to sexual violence, much of the discussion hinges on consent. MacKinnon (1989, p. 181-183) describes coercion and consent being evaluated as a "...socially reasonable, meaning objective man's, point of view..." which results in legal definitions based largely on rape myths. Similarly, more than twenty years apart Ehrlich and Tranchese (2001, p. 38, 2023, p. 23) discuss the persistence of the myth that 'real rape' is stranger rape, despite evidence from reporting showing that predominantly perpetrators are known to their victims (Office for National Statistics, 2025b, p. Main Points). Sexual violence within an existing relationship complicates consent because of possible coercion or threats. MacKinnon (1989, p. 175) suggests that the law divides women into two categories, the virtuous who are assumed to be un-consenting, and "Unvirtuous women like wives and prostitutes" who are assumed consenting by default, and that this explains why the onus is on the plaintiff to prove non-consent. Writing on the extreme case of sex trafficking, Munro (2008, p. 243) describes rhetorical use of deserving "ideal" victims who are invoked as a benchmark against which to evaluate actual victims. Usually found wanting, her actual trafficking victims were not just disbelieved but criminalised as "complicit agents". Beyond proving non-consent, evidence of resistance may be required with an expectation of fight or flight narratives, rather than the less discussed but more common freeze or fawn responses to trauma (Yapp, 2025, pp. 2, 6). Thus developments in our understanding of normal responses to traumatic events may not be utilised, with normal victim behaviour reprocessed by the defence team as proof of unreliability.

Cross examiners' questions can directly call on or imply rape myths reflecting a "sexist and androcentric belief-system informing and authorising" sexual violence, potentially influencing juries and court professionals (Ehrlich, 2001, p. 63). Such narratives and arguments exert a powerful force on jurors when combined with their own socially constructed beliefs (Coulthard & Johnson, 2007, p. 101). Complainants ability to tell their own story is limited because lawyers control the topic, and ask questions that are "highly constrained and constraining", which can position a witness as unreliable (Coulthard & Johnson, 2007, p. 106). Ehrlich (2001, p. 62) detailing the analysis of two cases against a single perpetrator, found that claimants were frustrated in their attempts to tell their own

stories, restricted by the cross-examiner's questions. She went on to identify extensive manipulation by legal professionals who were empowered to "present, characterize, limit and otherwise direct the flow of testimony" effectively reframing the claimant's story. This approximates to Heffer's (2015, p. 7) term "emplotment" (constructing the "right" story to fit the charges, but also to connect with the jury's mental scripts). Thus defence arguments incline towards binary understandings of consent associated with violent stranger rape, rather than nuanced understandings involving fear or coercion associated with the more common acquaintance rape. Erlich (2001, p. 148) concludes that by rationalising actions with reference to rape myths, a perpetrator and legal professionals may maintain "culturally-sanctioned ignorance" that avoids recognition of the issue and having to deal with it.

Reflecting social progress, FLT has expanded to include many other fields such as motherhood and family, un-paid care work, intersectionality, technology and social media, LGBTQ+ issues, and reproductive justice. Some of these will now be discussed.

Since the 1960s social changes have led to a growing variety of types of family units, however law and policy remain oriented towards the traditional nuclear family (Fineman, 1995, p. 4). Fineman (1995, pp. 9, 162) calls for care-taking of others and dependency to be central within the law's definition because everyone is dependent on the care of others at some point in their life. Fineman's re-definition of family thus extends to all configurations of caregiver(s) and dependant(s). Her perspective highlights that care-giving is gendered with the burden more often falling on women, and within her proposals she calls for a "refashioning" of masculinity and fatherhood (Fineman, 1995, pp. 204, 205).

In addition to barriers faced by white women, Trinch (2003) describes the intersectional cross-cultural challenges faced by immigrant Latino domestic abuse victims trying to access the US legal system. Generally having English as a second language led to the content and delivery of women's responses to questions being misinterpreted by their non-Latino interviewers (Trinch, 2003, p. 46). Complainants' motives were evaluated against US ideological standards (Trinch, 2003, p. 16), and their emotional responses to questions were misconstrued by interviewers as "flighty" and "irrational" or "cold, calculating and conniving" (Trinch, 2003, p. 178) damaging their prospects of success.

Concerns of inequalities within the legal system have been broadened beyond racial and ethnic groups to include issues such as "sexuality, transgender identity, and masculinity" (Conaghan, 2013, p. 74). With greater awareness of the complexity and fluidity of gender categories Conaghan (2013, p. 76) offers a nuanced view of which group might "probably" benefit from inequalities "white, middle-class, heterosexual, able-bodied men". This is the demographic most closely associated with professional law careers in the US and UK, where men have better career prospects than women who still experience marginalisation (Conaghan, 2013, p. 79) suggesting that the situation is likely to continue. Conaghan (2013, p. 82) identifies the failure of formal workplace equality to deliver the hoped for equality in practice because the system's values and its measures are male-centric. An example would be labour laws that position pregnancy as atypical and requiring "special pleading" rather than being a normal attribute (Conaghan, 2013, p. 82).

Social media has represented a form of activism outside of the legal system, that by-passes the editorial process of news. Those who are victimised but lack a voice, and therefore

denied justice through the legal system, have used this opportunity to be heard, an example of which is the #MeToo movement on Twitter (Tranchese, 2023, pp. 329, 330, 344, 402). While #MeToo has enabled victims to provide each other with support, it has also resulted in a “monumental” backlash against accusers by those seeking to silence the victimised and protect perpetrators (Tranchese, 2023, p. 347).

The following three sections outline some theories suggesting how basic human traits may operate to maintain society’s beliefs and values including those underpinning inequality and discrimination.

#### **2.4.4 Just World Theory**

It is not unusual for third parties to disbelieve victim’s claims of sexual violence, often by invoking rape myths<sup>19</sup> that have been discredited for many decades. Often victims seem to receive unduly harsh treatment within Twitter. A possible explanation was provided by Lerner and Miller (1978, p. 1031) who carried out experiments inflicting simulated electric shocks on pseudo-victims, similar to Stanley Milgram’s (1963, p. 372) famous experiments into authoritarianism. Lerner and Miller’s resulting Just World Theory (JWT) describes a protective measure that enables people to believe that life is fair and bad things don’t happen to ‘good’ people like themselves. To maintain their worldview it is necessary to believe that if something bad has happened to someone then they must deserve it, or in some way have brought it upon themselves. The uncomfortable alternative is that bad things can happen randomly to good people doing the right things and therefore the world is unfair. In an unfair world people cannot protect themselves from bad things, leaving them in a constant state of vulnerability. Victims themselves may even adopt this view retrospectively, rationalising that if the bad experience was their fault, they can do something differently next time and it won’t happen again. The (debunked) rape myths provide a mechanism for restoring JWT and the perception of safety. Here the word *world* in JWT implies a personal world rather than a global context, “As events become closer to their world, however, the concern over injustices increases greatly, as does the need to explain or make sense of the events” (Lerner & Miller, 1978, p. 1031). Hence, those who identify most closely with the victim may most strongly subscribe to JWT and blame most adamantly.

Fundamental Attribution Theory (FAT) (Ross, 1977, p. 184) is similar but goes further in explaining bias in individual’s attribution of motives to others’ actions. FAT suggests that people are inclined to overestimate the part played by others’ individual traits and underestimate external environmental factors. This aligns with the questioning of victims’ character and personal habits, and chimes with Ross’s observation that people have a “deep conviction that personal dispositions control and are reflected in everyday social behavior” (1977, p. 187).

#### **2.4.5 Reward Expectation Theory**

Reward Expectation Theory (RET) can relate to pay, evaluation, and promotion in the workplace (Wagner & Berger, 1997, pp. 9, 12). RET posits that high status individuals are judged as being more skilled and higher performing. As others react to them with positive

---

<sup>19</sup> Appendix A – Rape myths

evaluations their confidence increases as does their certainty and assertiveness. In contrast, lower status individuals provide more positive reactions to higher status individuals and allow them more opportunities to speak, suggesting a mutually reinforcing and polarising effect. Wagner and Berger also found that level of pay affects self-evaluations of capability, so when someone is underpaid they may internalise this as evidence of their performance being weaker (1997, p. 16). If higher valued (and higher paid) positions are predominantly held by men this reinforces the expectation and acceptance of a power, prestige and pay differential between male and female employees (Wagner & Berger, 1997, p. 20). Consequently, gender can become a proxy for power. In a mixed gender work team a woman behaving assertively and against the expected gender stereotype may be perceived as lacking legitimacy and therefore resisted. This theory, proposed twenty-two years after equal pay legislation (*The Equal Pay Act 1970*, 1970) was enacted in the UK, describes an unconscious but resilient mechanism for the continued undervaluing and underpayment of women's skills. It also contributes to the barrier women face in negotiating pay rises, discussed in section 6.4.2.2, and 6.5.6.

Two decades later a lab-based equal pay study by Auspurg et al (2017, p. 195) found that when both women and men were asked directly they supported gender pay parity. However, when questioned on specific job-based vignettes, both sexes to an equal degree evaluated women as deserving only 92% of the pay ascribed to equally qualified men suggesting underlying gendered beliefs (Auspurg et al., 2017, pp. 202-204). Similarly Cooke (2024, p. 358) found that even when pay levels were attributed arbitrarily, those assigned greater economic resource were perceived as having higher status and power. Furthermore, this was seen as fair by both the higher and lower paid party. Cooke also noted that the performance of male employees was perceived as driving earnings while performance did not influence female employees' pay. A self-reinforcing relationship was noted between under-payment and perceived unreliability and lack of commitment of women to their jobs. Stereotyped expectations such as women deprioritising their career in favour of caring responsibilities thus become a self-fulfilling prophecy. This highlights that when parental leave is taken predominantly by women it further embeds the view that women are less committed to their jobs than their male colleagues. Forty-two years after the Equal Pay Act was enacted in the UK (54 years in the US) Auspurg et al note:

*“This deeply internalized acceptance of gender differences and inequality likely makes gender status beliefs strong, self-fulfilling, and thus stable over a long period of time. (Auspurg et al., 2017, p. 203).*

RET's premise that rewards are more driven by perception than performance suggests a shift from conscious gender discrimination which has been seen and legislated for, towards deep seated and largely unconscious discriminatory values.

#### **2.4.6 Labelling Theory**

The power of language makes the words used to talk about individuals, and how people refer to themselves, very important. Labelling theory suggests that labelling and self-labelling has the power to influence identity and drive behaviour. Furthermore, any stigmatisation may “function as a means of formal social control” or “as a means of removing these minorities from various avenues of competition” (Goffman, 1990, pp. 164,

165). Goffman suggests that individuals are labelled as “deviants” for disaffiliating from “the social place accorded to them”<sup>20</sup> representing “failures in the motivational schemes of society” (1990, pp. 170, 171) and rocking the boat. This is important because the stigma of a powerfully discrediting label may be internalised, shaping the individual’s self-concept and social identity (Goffman, 1990, pp. 13, 19). Re-situating Labelling Theory within social media with its hate speech, trolling, and capability to go viral, makes a powerful and potentially damaging combination.

Other more recent work has focused specifically on the use of the VICTIM and SURVIVOR labels as applied to the subjects of sexual violence. For example Williamson and Serna (2018, pp. 669, 673) found that many of those experiencing sexual violence preferred the label SURVIVOR for its perceived agentic connotations rather than VICTIM which is used by the police and legal systems. However, of their 85 participants, while 35 referred to themselves using SURVIVOR, 24 chose VICTIM, and 26 thought of themselves by neither label, suggesting a variety of participant interpretations.

Messamore and Paxton’s (2021, pp. 13, 14) study of 3,756 nonprofit organisations found that those that referenced women in their literature increased their use of SURVIVOR alongside a slower rate of decrease in the use of VICTIM. Greatest use of VICTIM and SURVIVOR was found within organisations that focused on sexual violence. Non-profits that did not specifically address women had lower usage of both VICTIM and SURVIVOR. They concluded that both terms were gendered and may refer to a “single narrative through which women are seen” (Messamore & Paxton, 2021, p. 16). In addition to being damaging to women, this is likely to cause difficulties for men who are victimised, and this would make a useful topic for another study.

## **2.5 Victims, survivors and the silenced**

This section explores some literature concerning the terms VICTIM and SURVIVOR. In Chapter 1 both the traditional and updated definitions given for VICTIM and SURVIVOR are quite different, however the everyday examples seen in the present study often imply more flexible usage. Reflecting this the Collins Dictionary entry for VICTIM gives SURVIVOR as an antonym. (n.d.-b). Cole alludes to this complexity describing the words VICTIM and SURVIVOR as “ensnared in ideological webs” (2021, p. 39) each defined as the opposite of the other. Existing research in this area comes from the disciplines of: Criminology; Victimology; Law; Journalism; Healthcare; Psychology; and Feminism. Some of this work will now be discussed to provide background and provide support for Chapters 4 and 7.

The Ideal Victim stereotype is described in section 2.5.1, while the relationship between VICTIM and legal system is explained in section 2.5.2. Section 2.5.3 discussed representations of the victimised in news and social media, while in section 2.5.4 the denigration of VICTIM is contrast with the valorisation of SURVIVOR in 2.5.5, and the potential obligation of survivors is explained in section 2.5.6.

---

<sup>20</sup> Suggesting the observation (Ammerman & Groysberg, 2017, pp. 1, 2 (in pdf reader)) in section 2.2.3 that sexual violence is used as a means of “keeping women in their place”.

### 2.5.1 Ideal Victim

In Christie's 1986 book, reproduced in Sloan et al (2018, pp. 11-22), he introduced the term Ideal Victim to describe a person universally viewed as legitimately deserving of sympathy. The Ideal Victim is referred to throughout the present study particularly Chapter 7. This stereotype is characterised as: weak; engaged in a creditable project; devoid of blame; and having been wronged by a large, malign and unknown offender. This describes an individual in a specific situation. It is easier for a child to qualify, difficult for a woman, and more problematic still for a man. Particular difficulties arise from the inclusion of the *weak* aspect which confounds any adult who wishes to be seen as an agentic and self-determining individual. In claiming structural discrimination, it necessarily positions victims as weak and vulnerable (Jones et al., 2022, p. 9). A second difficulty is that sexual violence perpetrators are commonly known to their victims. Christie notes that the closer women get to achieving full equality the more difficult it will be for them to achieve Ideal Victim status (Sloan et al., 2018, pp. 20, 21).

Public debate surrounding the VICTIM label is often emotive because "claims to victimhood are claims to power" (Chouliaraki, 2024, p. 16) immediately setting the individual at odds with the "weakness" of the Ideal Victim stereotype. Chouliaraki found that the debate around victimhood intensified following the 2008 financial crisis and the 2019 covid19 pandemic because inequalities widened with the poor, working and middle classes feeling angry and revengeful (2024, pp. 19, 20).

Cole (2021, pp. 31, 32) suggested that the release of the film Shoah (Lanzmann, 1985), raised the profile of the holocaust in the US mainstream media, and gave the word VICTIM a heightened moral authority. She went on to suggest that this set the suffering of the Jews as a new measure of victimhood and moral authority, raising the bar to an almost unattainable height for other situations. Thus the term VICTIM becomes a claim for moral authority and power making its use by bogus victims appear even more reprehensible. Cole (2007, pp. 6, 175) uses the term *anti-victimists* (analogous to *racists*) to describe those who attack *victimists* (bogus, or undeserving victims). Anti-victimists are characterised by their view that the truest of Ideal Victims are those punished by political correctness, affirmative action and "injurious victim politics" (Cole, 2007, p. 6). She suggests that it is an identifying mark of anti-victimists that after denouncing a victim as bogus they go on to declare themselves the True Victim<sup>21</sup>. This study uses the term *bogus victim* to avoid confusion with many other evaluative adjectives analysed in later chapters. When trying to identify the bogus victim in a dispute Chouliaraki (2024, p. 36) advises asking the questions "Which positions of power do these selves speak from? What kinds of benefits do their claims accrue to them and the communities they bring together?"

Returning to Cole's work on the film Shoah, she argues that its holocaust testimonies elevated and created new positive meanings for the label SURVIVOR (2021, pp. 28, 32). This

---

<sup>21</sup> Cole refers here to a slew of books published in the 1990's identifying various True Victims: Charles Sykes, A Nation of Victims: The Decay of the American Character (1992); Robert Hughes, Culture of Complaint: The Fraying of America (1993); Shelby Steele, The Content of Our Character: A new Vision of Race in America (1990); Alan Dershowitz, The Abuse Excuse: And Other Cop-Outs, Sob Stories, and Evasions of Responsibility (1994); and Dinesh D'Souza, Illiberal Education: The Politics of Race and Sex on Campus (1991).

included an association with agency through taking responsibility and persevering to overcome a personal experience and gain greater social acceptability (Cole, 2007, p. 138). Significantly, in claiming survivorship the experience itself and any harmful consequences are shifted into the past, potentially preventing damage from being dealt with.

### **2.5.1.1 Ignorant victims**

In addition to those victims who do not qualify as ideal in section 2.5.1, Christie also describes a second group that he calls “ignorant victims”, that is those who are either: unaware of being victimised; or are unaware of who is victimising them (Christie, 1986, pp. 17-18). His example is that of a socially deprived school child who is unaware that he is competing with more advantaged others (Christie, 1986, p. 12). The child interprets his poor results as a personal failing rather than resulting from a system designed by another class for their own children. Christie’s point here is that with no visible perpetrator it becomes difficult for individuals to see themselves as victims, more difficult to be recognised by others as victims, and more difficult to be seen as sufficiently disadvantaged to obtain justice (Christie, 1986, pp. 17-18). Christie identifies two contributory factors, firstly the “contrary interests” of others, and secondly a cultural view of barriers as personal responsibilities alongside an obscured or faceless perpetrator.

### **2.5.2 VICTIM and the legal system**

An understanding of the differing implications of the labels VICTIM or SURVIVOR is fundamental to this study. The “United Nations Human Rights office of the High Commissioner Declaration resolution 40/34” defines what “Justice for Victims of Crime and Abuse of Power” looks like (1985). The word SURVIVOR does not appear in this document and there is no parallel document for survivors, clearly and squarely associating VICTIM, but not SURVIVOR, with justice. This is reflected within the UK and US criminal justice systems where crimes are discussed exclusively as being perpetrated against VICTIMS. The legal profession is likely to work to maintain their established and legally proven terms resisting changes occurring in the wider society (Sinclair & Carter, 2004, p. 151). At the same time, there is some evidence of certain aspects of legal services acknowledging the sensitivity surrounding sexual violence. In the UK for example the Sentencing Council’s report “Attitudes to Sentencing Sexual Offences” (McNaughton Nicholls et al., 2012) uses the combined *victim/survivor* allowing for either preference, and the US Office of Justice Programs sponsored “Victim or Survivor: Terminology from Investigation Through Prosecution” project (SAKI, 2021) opts for the hyphenated *victims-survivors* more suggestive of a continuum. While these solutions provide flexibility they also allude to lack of agreement around the meanings and perceptions involved.

The 2021 UK report “The end-to-end rape review report on findings and action” found that “too many rape victims do not receive the justice they deserve” (H M Government, p. 3). This report recommended that victims should receive more protection against intrusive and irrelevant questioning on personal, sexual or therapy histories, and all legal practitioners are trained in misconceptions and rape myths (H M Government, 2021, pp. 33, 47). Disappointingly, in 2025 the Law Commission found that despite some progress, “the criminal justice process for rape and serious sexual offences is still flawed” with particular concerns including the perpetuation of rape myths (Law Commission, 2025). Two decades

earlier Smart foresaw that while action must be taken to improve laws, "... we should not assume that the law can provide the solution to the oppression that it celebrates and sustains." (2002, pp. 53, 55).

Over the last few decades some countries such as the US, Canada, Australia, and New Zealand have implemented Rape Shield laws to protect accusers from intrusive and irrelevant interrogation of their personal and sex lives, often rooted in rape myth reasoning (Gröndal et al., 2025, pp. Discussion, 3rd paragraph). In the UK Rape Shield protections rely upon the European Convention on Human Rights (Council of Europe, n.d.). However, these legislative protections do not appear to have resulted in better outcomes for victims (Cowan, 2019, pp. 22, 23). Over time as rates of reporting of sexual violence have increased across Europe, corresponding rates of both prosecutions and convictions have fallen, "... reminding law and policy-makers that law reform, in itself, does not equate with better outcomes, if by better outcomes we mean more convictions" (Cowan, 2019, pp. 28, 33, 34). Smart (1989, pp. 10-13) views the law as aligned with a powerful "masculine culture" associated with a "malevolence" towards women's concerns, thus situating any challenge to the law as taking place within a gendered cultural setting. This may contribute to the persistence of rape myths within the legal professions, judiciary, and juries (Cowan, 2019, pp. 34-36) and also police officers (Gröndal et al., 2025, p. 1133), creating erroneous expectations of victim behaviour that contradicts victims experience and prevents them from fully conveying their narrative.

A strongly resilient rape myth is that of the violent stranger rape. Evidence from the 'Crime Survey for England and Wales' (Ministry of Justice, 2018, p. 6) reported that for their "Most Serious" offences category 90% of perpetrators were known to their victims, and 56% were partners/ex-partners. Importantly, this situates victims as commonly embedded in social relationships with their assailant. This is at odds with the archetypical but less common "real rape" by a violent stranger, which may lead police or legal questioning to focus on a victim's ongoing communications or contact with a perpetrator, and consequently conclude that sex was consensual. Victim behaviour that contradicts expectations, for example appearing calm after the assault, may be interpreted as an abnormal response or mental health issue, which is then used to undermine the integrity of the victim's memories and testimony (Yapp, 2025, p. 2). Thus the victim's credibility is damaged and their claim that they did not consent seen as unreliable.

The rape myth expectation that "real" victims would fight off an attacker or flee, may leave those that do neither as presumed to have consented to sex (Smart, 1989, pp. 39, 42). For example, submission to sex in fear of violence or losing their job may be misconstrued as consent and an agentic choice to remain (Bailey et al., 2023, p. 5). However, research indicates that of the flee, fight, freeze, or fawn autonomic fear responses, the less familiar freezing and fawning are more likely (Yapp, 2025, p. 6). Furthermore, the transitory freeze response can progress into Tonic Immobility (TI), in which trauma induced paralysis renders any physical or verbal resistance impossible, which may also be misinterpreted as passive consent (Moller et al., 2017, p. 1 (in pdf reader)). Freezing is an adaptive response to a life-threatening situation that may calm the perpetrator and reduce the risk to the victim (Bailey et al., 2023, pp. 2, 3). Möller (2017, p. 7 (in pdf reader)) found that 70% of her rape victim participants reported symptoms of TI, and their likelihood of experiencing TI was doubled

for: victims of previous sexual assaults; those experiencing more severe assaults; and those with pre-existing PTSD. Unaware of why an accuser might not fight back or run may result in both victim-blaming, and the victim self-blaming.

When fight or flight are not possible, victims may try to mitigate against aggression or retaliation by fawning, defined as “overly helpful towards threatening, unpredictable, or invalidating authority figures” (Schlote, 2023, p. 2). Fawning is a survival mechanism whereby the victim unconsciously responds to “exploitation, shame, neglect, abuse, or other harm”, with behaviours that may calm their assailant and reduce risk (Clayton, 2025, pp. 19, 21). As a clinical psychologist, Clayton found her clients could continue to fawn after the event, due to an ongoing state of fear of future harm to themselves or others (Clayton, 2025, pp. 21, 24). This is particularly relevant when there is a social or work relationship with the perpetrator, running the risk that those unaware of this phenomena may interpret a victim’s fawning as proof of consent rather than fear. Lack of understanding of freezing or fawning responses may hinder the healing of victims through self-blame, and result in inappropriate questioning and evaluation within the court room. The language used in court cases is important because it is often quoted in reporting of sexual violence crimes, for example phrased as the victim *telling* police, jurors or court, and this context shapes how the victim can tell their story (Tranchese, 2023, p. 383). Representations of the victimised in news and social media are discussed in the next section.

### **2.5.3 The victimised in news and social media texts**

News articles are selected for publication because of their potential interest to readers, and because they reflect the political leaning of its editors and target audience (Semino, 2018, pp. 378, 379) making them relevant to a linguistic study. The relationship between traditional news outlets and social media is complex, with most commercial news broadcasters now also operating on social media, with journalists likely to have their own social media accounts. Within Twitter sexual violence coverage by traditional news articles is often discussed, with links to the original article, these tweets may be retweeted many times. Conversely news articles may discuss online activity such as the #MeToo hashtag.

In a study based on high profile news coverage of the Steubenville rape case, Barca (2018, p. 279) found that “negative agency as a grammatical expression of victim-blaming”<sup>22</sup> manifested as: episodic framing<sup>23</sup>; foregrounding the perpetrator’s perspective; positive and passivated perpetrator characterisations; and negative and agentic victim characterisations. Barca suggested that journalists were blatantly doing some of the work for the legal defence teams (Barca, 2018, p. 279). She found that journalists were framing societal problems as personal decisions and responsibility, suggesting “...neoliberal notions of risk that ignore structural gender inequality as a cause of rape.” (Barca, 2018, p. 276). Reviewing studies across different types of news media in the UK and US, she concluded that crimes were portrayed in isolation rather than as part of a broader pattern, because events that are

---

<sup>22</sup> Barca explains “Negative agency” as a neoliberal concept that attributes agency associated with a negative evaluation to an individual, suggesting they brought the outcome upon themselves.

<sup>23</sup> “Episodic framing” refers to the representation of incidents as isolated events rather than as part of a broader or systemic pattern.

common are not newsworthy (Barca, 2018, p. 276). Within the Steubenville reporting she also found “episodic framing and a view of risk-taking consistent with neoliberal victim theory” (Barca, 2018, p. 277), leading her to suggest that the journalists’ worked within power structures reflecting “neoliberal and patriarchal appeasement” (Barca, 2018, p. 280). Attributing neoliberal agency to victims represents them as culpable for their own victimisation (Barca, 2018, p. 266) often visible as rape myths in news stories. News media are commercial businesses that may favour products that prioritise sales and revenue by publishing articles quickly, potentially risking an over-use of clichés, and language reflecting potentially erroneous cultural assumptions (Barca, 2018, p. 268).

Through the twenty tens a rise in news reporting of celebrity sexual violence cases coincided with an increase in hedging using the term ALLEGED INCIDENT, suggesting growing disbelief of victims alongside a reluctance to risk defamation suits from celebrity defendants (Tranchese, 2023, p. 217, 218). Moreover the choice of stance adverbs used in describing the accused as “vigorously” or “categorically” denying, suggesting “outrage, indignation, and affront”, positioned victims’ claims as ludicrous or unreliable (Tranchese, 2023, p. 191, 192).

### **2.5.3.1 Social media**

Social media is globally influential and important, enabling public figures to interact directly with their audience. Twitter particularly is an important forum for news content, making it a necessary part of professional life for many women and other minorities (Amnesty International, 2018, p. Chapter 1). Amnesty International considers that online violence contravenes the human right to freedom of opinion and expression, and constitutes gendered abuse. Victims of sexual violence were speaking out long before #MeToo, but were often ignored by traditional news providers (Tranchese, 2023, p. 402). With social media, women gained direct access to a “megaphone” that bypassed editorial processes to call out sexual violence. Furthermore, #MeToo broadened the scope of sexual violence to include previously “normalised practices (most notably sexual harassment), which we are socialised into accepting as a normal part of life” (Tranchese, 2023, p. 338).

Twitter is not a neutral space. In creating communities the excluded Other, who is not a member of the community, is also created. Social media platforms tailor content to users’ existing preferences and affiliations using algorithms, creating echo chambers or nodes of similar content, tending to confirm existing views and distancing alternative opinions and evidence (Bouvier, 2022, pp. 182, 192). She went on to suggest that the #MeToo hashtag was not driven by attempts to identify and solve problems, but carried by affective “expressions of engagement, solidarity, moral indignancy and righteousness” (Bouvier, 2022, p. 193). This supports and encourages simple good/bad thinking that is required for successful social media movements (Bouvier, 2022, p. 188). Moreover, rather than reflecting their beliefs, account-holders may use algorithms to deliberately self-promote by posting high energy negative tweets likely to increase their traffic (Bouvier, 2022, p. 189).

Zappavigna (2012a, p. 4 (in pdf reader)) introduced the term *ambient affiliation* to describe how Twitter hashtags like #MeToo invite users to “Search for me and affiliate with my value”. Thus hashtags, such as #MeToo, enable individuals to form bonds, make communities, compare and rework shared values around current topics. The more searchable utterances are, the greater the potential for bonding (Zappavigna, 2012b, p. 194) creating a sense of

belonging without actual direct interaction (Zappavigna, 2017, p. 216). While hashtags can be used to provide support and solidarity for accusers through ambient affiliation they are public, so can also be used to challenge the originators of the hashtag, as described by Jones et al (2022, p. 6) in relation to #NotAllMen. #NotAllMen was a defensive response to the furor following the rape and murder of Sarah Everard, intended to highlight that not all men were rapists and murderers. However, this was seen as misogynistic for backgrounding the crimes committed to Sarah Everard and the pervasiveness of sexual violence against women. Consequently, #NotAllMen was also used to challenge misogynistic behaviour that “not all men confront when they see it” (Jones et al., 2022, p. 7). Victim-supportive and victim-critical hashtags are open to challenge, and where tweets degenerate into “hostility and gendered cyberhate” they could deter engagement with social media (Palomino Manjón, 2024, pp. 107, 108). Abusive responses are strongly associated with women expressing opinions about politics, current affairs, or feminism (Amnesty International, 2018, p. Chapter 5).

Trying to lessen abuse women may: self-censor their posts; filter or consciously not read incoming messages; or even remove themselves from Twitter, to the detriment of their professional lives. For example, 31% of women experiencing online abuse in the UK and 35% in the US no longer post on certain issues (Amnesty International, 2018, pp. Chapter 5, 6). For many people online life is real life now<sup>24</sup> (Amnesty International, 2018, p. Chapter 5) so online threats may involve real world harm. Moreover, public online abuse adds humiliation, increasing its potential to deter others from speaking out. The then First Minister of Scotland Nicola Sturgeon feared it may discourage women from entering politics (Amnesty International, 2018, p. Chapter 5).

Tranchese (2023, p. 35, 370) suggests that social media’s supposedly democratic equal access to obtain justice is an illusion. More often their business models favour posts that maximise revenue, rather than those promoting complex feminist histories around the gendered term VICTIM, or those discussing cultural barriers.” (Banet-Weiser, 2019, p. 174). The #MeToo hashtag is bound into Twitter’s business model which is geared towards individual consumers, rather than supporting the collective action required to counter a structural discrimination such as sexism (Banet-Weiser, 2019, pp. 173, 174).

The hashtag #MeToo has been referred to as though it were a person by accusing it of failing to bring about real world change and leaving behind victims who are not “rich, white and famous” (Gill & Rahman-Jones, 2020). However, this mistakes #MeToo for an organised political movement, confusing its role with that of policymakers, and ignoring structural barriers such as poverty (Bouvier, 2022, p. 185; Tranchese, 2023, pp. 336, 351). However, personalisation enables it to be held responsible and accused of not doing enough, and also of going too far (Tranchese, 2023, pp. 353, 356). By victim-blaming: the personalised hashtag; feminism; and women generally, attention is drawn away from the underlying structural causes (Tranchese, 2023, p. 443). Detached from historical feminism and the context of structural barriers, #MeToo became isolated, problematic, and potentially scapegoated (Tranchese, 2023, p. 461) with tweets often contentious.

---

<sup>24</sup> Vonny Moyes, journalist (Amnesty International, 2018, p. Chapter 5)

#### **2.5.4 The evaluation and debasement of VICTIM**

The following discussion of the negative prosody accumulating around VICTIM draws on work from Criminology, Journalism, Psychology, and Feminism. Starting with an illustrative opinion piece for *Harpers Bazaar*, journalist and rape victim Campoamor (2018, pp. 1, 2 (in pdf reader)) describes the punishing retaliation that can follow speaking out as a victim. Having been called: a “professional victim”; a “self-pitying victim”; and a “snowflake ... perpetuating a victim culture”, she observed that: “We have bastardized the word to the point that it’s used to diminish, discredit, and disparage anyone who has endured the worst of humanity”. Researchers suggest various reasons for this type of denigrating response.

Those victimised by sexual violence who fail to meet the high standards of Christie’s (1986, p. 12) Ideal Victim may be attacked as bogus victims who indulge in victim politics and positioned as posing a deliberate threat to the US welfare state and society itself (Cole, 2007, pp. 11, 23). Gradually the term VICTIM has become tarnished, used as a term of abuse and avoided even by those suffering dreadful crimes that clearly qualify as victimisation (Cole, 2007, pp. 22, 138). For example, the *New York Times* reported in 2020 the attack on Nicole Barrett in which a paving slab was used to cave-in her skull, but she later “insisted that she wanted people to know that she is not a victim or a ‘half-dead person.’” (Finkelstein, 2021).

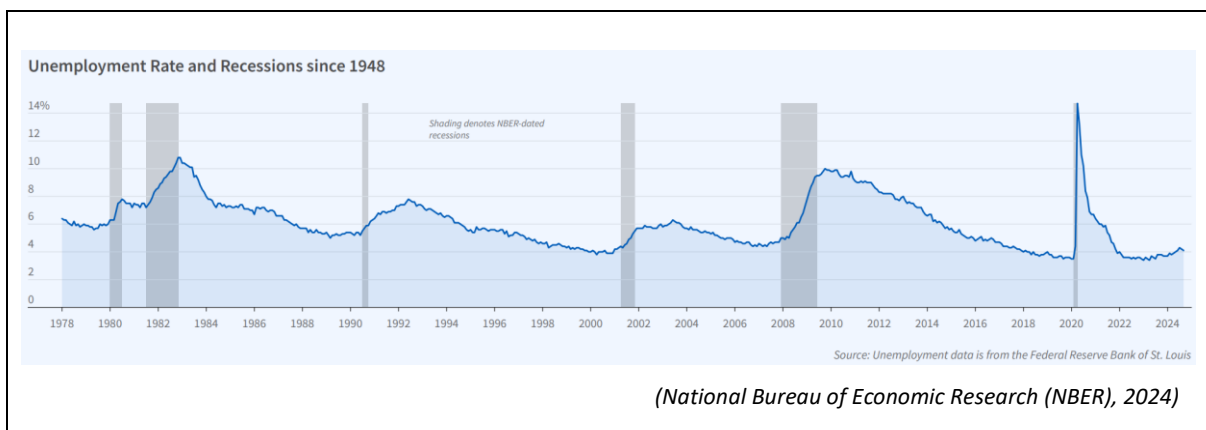
Exploring the associations around the word VICTIM, Norwegian criminologist Van Dijk (2009, p. 3) was an early voice advocating against the use of the term within the criminal justice system, and he provides some relevant background. His view was that the etymology of VICTIM in the suffering of Jesus Christ as the archetype of sacrifice invoked a response of both compassion for the victim but also an expectation of passivity and forgiveness of the perpetrator. He went on to suggest the consequence that the victims of crime are required to restore the peace disrupted by offenders by declining acts of revenge or legal action. Additionally, van Dijk’s victim narrative analysis leads to “In extreme cases, victims openly contesting their expected roles of passive sufferers, may ... become the target of scapegoating mechanisms and be re-construed as offenders or accomplices.” (Van Dijk, 2009, pp. 14, 25). Interestingly, he found that the factor triggering negative responses and suspicion was a victim’s demonstration of autonomy and an actively resilient coping style which acts against the weak and passive victim stereotype.

Cole provides another perspective, suggesting that the US civil rights and feminism movements of the 1960s and 70s provoked a conservative backlash later in the 70s and 80s against progressive politics and the welfare state (2007, pp. 10-12). In this climate, attempts by disadvantaged groups to organise against societal injustice were perceived by American neoliberalists as a threat (Mahoney, 1995, pp. 169-171). That threat was perceived as likely to result in an unfair burden on welfare resources at times when the nation’s economy was faltering during recession (Figure 6). Reducing the size of the welfare state, and the demand placed upon it, is part of Republican and neoliberal world views, discussed earlier in section 2.3.2. The period of data gathering for this study largely lies within Donald Trump’s (Republican and conservative) first term as president of the US, Jan 2017-Jan 2021. The political right views citizens’ as responsible for their personal safety and success in life in effect victim-blaming, which “facilitates and justifies class inequalities and, at the same time, is rooted in a collective unconscious” (Cole, 2007, p. 115).

This continues a historic tendency to repress those who challenge the status quo by portraying them as “a subversive threat to the stability of the nation” (Cole, 2007, p. 4). Cole goes on to describe the characteristic ways that critics seek to undermine victims by (Cole, 2007, p. 27) (examples appear in later chapters):

1. Claiming that many victims making accusations are bogus;
2. Claiming victims with credible claims are manipulating them to gain disproportionate benefit;
3. Even if they are “real” victims, claiming victimhood harms them and society;
4. Claiming that the individuals making accusations are doing so aggressively.

Figure 6, US recessions from 1980 – 2024



Mahoney (1995, p. 62) points out that as agency is a personal trait it can become a barrier to victims identifying each other, preventing mutual support and organisation, and reducing visibility of societal issues. Mahoney argues that as the concepts of victim and agency are diametrically opposed, if VICTIM is defined as one without agency then “Agency does not mean acting for oneself under conditions of oppression; it means *being without oppression*, either having ended oppression or never having experienced it at all” (1995, p. 64).

A number of Feminist, Psychology, and Sociology researchers provide useful background on the perpetuation of rape culture on Twitter. Stubbs-Richardson, Rader and Cosby (2018) analysed tweets immediately following the US Torrington and Steubenville Rape Trials, and Rehtaeh Parson’s rape and subsequent suicide. They found three main themes: “the virgin-whore binary and Just World Belief”<sup>25</sup>; “informational tweets as subnews”; and in support of victims “rape myth debunking” (Stubbs-Richardson et al., 2018, p. 98). They raised concerns about victim-blaming accounts and tweets having firstly, more followers; and secondly, were more frequently retweeted than those in support of victims (Stubbs-Richardson et al., 2018, pp. 100, 103, 105). This dynamic is designed into platform algorithms because they “monetise claims and counter-claims” (Alison Phipps, 2019, p. 68) potentially enabling victim-blaming tweets to spread quickly and widely online to tens of thousands of readers amplifying their damaging effects.

---

<sup>25</sup> Just World Belief (JWB) and Just World Theory (JWT) are terms favoured by different individuals to refer to the same thing, refer to section Just World Theory (section 2.4.4).

Removing the minority who are Ideal Victims leaves those who are present but discredited, and linguistic traces of others characterised as deserving, but who shun victimhood by remaining silent. The missing and silent are explored in section 2.5.7 while the discredited are discussed further in the Chapter 4.

### **2.5.5 SURVIVOR as active coping**

Having discussed the challenges associated with the VICTIM label in the previous section this section shifts attention to the active adoption of the SURVIVOR label, which contributes to the analysis in Chapters 4 and 7.

Social media platforms such as Twitter have given women the ability to tell their stories in their own unedited words in a way that was not possible before, but this has not prevented rape myths from continuing to circulate (Palmer-Mehta, 2018, p. 160). The 35 women drugged and raped by Bill Cosby have consistently portrayed themselves, and been portrayed by others, as SURVIVORS, some specifically rejecting the label VICTIM even when battling with the ongoing consequences. One of them, Green, is quoted as saying “People go, ‘Why haven’t you gotten over it?...’ suggesting that even these women are being measured against the heroic SURVIVOR archetype (Palmer-Mehta, 2018, p. 174). Palmer-Mehta also found that claiming the agency of SURVIVOR did not necessarily deliver the response expected. The women in her study found that the likelihood of family and friends believing them was reduced if they described any element of active agency, such as going to the perpetrator’s home, allowing the perpetrator to kiss them, or persuading a perpetrator to wear a condom (2018, p. 179 (note 177)). Furthermore, Mahoney (1995, pp. 62, 68) concluded that viewing the self as an agent could prevent an individual identifying their own experience as abuse. External pressure on the victimised to have “gotten over it” before they are able to marshal the necessary resources, may prevent processing of the crime and delay the healing process (Fohring, 2018, p. 160). Fohring speculated that the uptake of SURVIVOR by support services may be “actually depriving victims of the space to grieve their loss of self and Beliefs” (2018, pp. 154, 158).

A Health Sciences study (Rees, 2018, pp. 2-6) asked twenty young women living with a history of breast cancer about the meanings they had constructed around SURVIVOR. This qualitative exploration found that the majority of patients disliked the term and many were ambivalent. While cancer is an entirely different situation, the reasons given for declining the label may still be relevant and include that the ideal of SURVIVOR: feels too high to live up to; implies a serious or near-death experience; and denies ongoing difficulties. This silences emotions such as anger and grief and make it more difficult to ask for help.

A word of warning comes from Lakoff & Bucholtz (2004, p. 51) who argue that renaming an uncomfortable concept does not solve the problem. They put it “Since attitudes toward the original referent are not altered by a change of name, the new name itself takes on the adverse connotations, and a new euphemism must be found”.

### **2.5.6 Forced heroism of the survivor**

In her New York Times Magazine article “Forced Heroism of the ‘Survivor’”, literary critic Parul Sehgel (2016, p. 1 (in pdf reader)) places the shift in favour of SURVIVOR as originating in the child sexual abuse cases in the 1970s, before being rapidly taken up more widely. The

victimised, then adults, were actively choosing the term SURVIVOR over VICTIM because of its positive and agentic associations. Sehgel's article suggests that media culture now prescribes survivorship as almost compulsory. She points to fictional pop culture representations of powerful victims of sexual violence such as "The Girl with the Dragon Tattoo" and "Kill Bill", noting their contrasting portrayals, "the victim was abject, a figure of shame and isolation, the survivor is lithe and frequently well armed." (Figure 7).

Sehgel concluded that the association of survivorship with heroism, toughness, and resilience puts responsibility for healing on the individual, depoliticising the causes (2016, p. 2 (in pdf reader)).

Kossew (2019, p. 182) also draws attention to the "immense pressure" to adopt the "compulsory hopefulness and optimism" of neoliberal survivorship that requires individuals to heal themselves quickly and return society

to its previous productive and economic position (Favaro & Gill, 2019, p. 155). When published in news and magazine content, neoliberal survivorship can suggest a clean linear and uni-directional transition from *tragic* to *inspirational* that refocuses attention from the victim's trauma to its impact on the audience, distancing the victimised and the victimisation (De Sacco, 2019, pp. 195-197).

The use, both by the media and in everyday conversation, of the terms SURVIVOR and VICTIM to describe those victimised by sexual violence crimes has a strong effect on an audience's reading of the situation, and their consequent support for a wronged individual (Papendick et al., 2017, p. 19 (in pdf reader)). Specifically, Papendick et al found that the use of SURVIVOR could result in reduced social support. Interestingly, where SURVIVOR was used with no other information about an assault they found the audience estimated the crime as more serious suggesting SURVIVOR's traditional near-death meaning. Forced heroism is discussed again in section 7.4.1.

### 2.5.7 The missing and the muted victims

One of the biggest challenges of this study has been how to analyse the missing and muted acknowledged by Duguid and Partington as "both a paradoxical and hopelessly wide-ranging endeavour." (Marchi & Taylor, 2018, p. 39). They define *absence* as "... discourses, information or roles which might be expected to be present but which are not". This study uses *missing* for its greater sense of loss (Oxford English Dictionary [electronic resource], n.d.-t) to mean the same as Duguid and Partington's *absence*.

The invisible wronged groups of the gender pay gap are believed to approximate to the *ignorant victims* previously described in section 2.5.1.1, and also those with RET beliefs discussed in section 2.4.5.

Figure 7, Uma Thurman, the victim/survivor in the movie Kill Bill  
(The Quentin Tarantino Archives, 2003)



The silence of rape victims is “emblematic of powerlessness in our society” according to Ahrens (2006, pp. 1, (in pdf reader)). Silence about a wrong conceals the causes and consequences and thus prevents social change (Ahrens, 2006, p. 11 (in pdf reader)). Ahrens found that her survivors who had received a negative reaction were more likely to internalise rape myths causing them to feel guilty and consequently self-silence. Additionally, audiences may reflect cultural stigmas leading them to view disclosures of sexual harassment more negatively than disclosures of other types of trauma (Delker et al., 2020).

### **2.5.8 Which victims may speak, and who are believed**

In Foucault’s “The Discourse on Language and Truth and Power” reproduced in Medina and Woods (2008, p. 316) he proposes that prohibitions operate to control who may speak about what, and when and where they may say it. Of relevance to this study is that he continues that in the context of politics and sex, society’s prohibitions “...hold some of their more awesome powers”. In some cultures this is more overt than others but those who do not qualify as Ideal Victims, and are relatively powerless, may feel exposed to greater risk by speaking out.

In addition to being able to speak out, it is also important for victims to be believed. This becomes problematic when discourses on sexual violence continue to refer to long discredited rape myths (Crown Prosecution Service, 2021)<sup>26</sup>. Tranchese (2023, p. 19) suggests that two particular rape myths underpin all of the others: that rape is always carried out by strangers; and women lie about rape. Belief in the former leads to the conclusion that if the accused was not a stranger then most likely it wasn’t a crime. However, in England and Wales<sup>27</sup> 48% of rapes and 32% of “other sexual offences” were perpetrated by intimate partners (Office for National Statistics, 2025b, p. Main points). This risks the accuser being disbelieved, and distances the crime and perpetrator. Belief in the second myth also casts doubt on accusers’ veracity, shifting the balance of credibility towards a rapist’s denial over the accuser’s claim, although often couched as fairness, or the rational sounding “reasonable doubt”. This creates a sense of even handedness between accuser and accused which, which may not reflect the reality of the power relationship. Where consent is the issue, the rape myth “women lie” raises the possibility that “no” means “yes” (Tranchese, 2023, p. 19) leaving women disbelieved with no means of refusal. The undermining of the victim’s credibility is more insidious than direct victim-blaming, and may be more difficult to identify and challenge (Tranchese, 2023, p. 176).

Crimes, victims, or perpetrators may be represented as *alleged*, or unproven. *Alleged perpetrator* suggests that a crime was committed while casting doubt around the defendant who is innocent until proven guilty. In contrast, the *alleged victim* introduces doubt over whether a crime was committed at all, and the accuser may not be believed until proven innocent (Tranchese, 2023, p. 198). This potentially discredits the many victims who withdraw cases, or whose cases end in an “insufficient evidence” verdict. Tranchese (2023, p.

---

<sup>26</sup> provided as Appendix A – Rape Myths

<sup>27</sup> For the year ending March 2025

255, 378) noted that within her news media corpora the trend was of decreasing overt victim-blaming alongside more common subtle insinuations.

The importance of who speaks and who is believed was highlighted in 2024 when it was reported that Mohamed Al Fayed, the late owner of the Harrods store in London, was accused of sexual misconduct, but of 21 historic accusations received the Metropolitan Police had passed only two to the Crown Prosecution Service.

In Figure 8 a BBC article (De Simone, 2024) states that unspecified “serious questions” have been raised and one of these may be whether the police saw the victims as having the authority to raise a successful speech act against such a rich and powerful individual. Speech act failure

Austin (1975, pp. 98, 105) uses “illocutionary act” to describe the speaker’s use of language for an intended purpose, as distinct from “the perlocutionary act which is the achieving of certain effects by saying something”. This highlights that intention is no guarantee of a speech act being correctly understood.

Figure 8, BBC reporting of Al Fayed accusations



A less dramatic but more familiar situation is described in Solnit’s<sup>28</sup> (2019, pp. 2-4) popular book which recounts her experience at a party where the host continually tried to recommend that she read a book, despite her repeated attempts to tell him that she was the author of that book. The host failed to interpret her message, and her illocutionary act failed in effect rendering her invisible as the owner of her work.

Maitra and McGowan (2010, pp. 167, 168) identify two mechanisms that cause speech acts to fail both of which have been exacerbated by pornography. Firstly, they describe a convention of *no* literally meaning *yes* so the utterance *no* is changed to mean its opposite and the victim is silenced. The second meaning is a failure of communication because the listener mistakes the speaker’s intention believing the speaker is only pretending to refuse in line with the “playing hard to get” trope. Dangerously, these speech act failures remove the language with which to object or refuse.

<sup>28</sup> Rebecca Solnit is a US writer and an activist on a variety of topics including feminism.

### 2.5.9 Recognising victimisation as crime

Until an individual realises that an unwanted behaviour constitutes sexual violence or that they are subject to unfair working practices they have no grounds on which to speak out, so are silenced. This issue is explored in the analysis in Chapters 4 and 7.

A workplace-based study by Nielsen et al (2024, p. 1) interviewed Danish (predominantly female) employees who had experienced workplace sexual harassment. They found several reasons driving under-reporting including: difficulties defining what sexual harassment was; fear of career-damaging retaliation upon speaking out; and fear of a potential negative reaction within their local culture (2024, pp. 5-7). Nielsen et al suggest a feedback loop with cultural influences creating a climate of tolerance around harassment and tolerance of harassment normalising and reinforcing cultural values of what behaviours female employees should accept. One of the interviewees explained:

*"I have downplayed it because I thought it was normal. Because I have heard it from so many of my friends (...) I think – and that is the dangerous part – that it gets trifled because it happens to so many, because then there is nothing to complain about. That is just the way things are. (Dakota)"*

*(Nielsen et al., 2024, p. 7)*



Fear of consequences is not without justification. Some of their interviewees who had called out unacceptable behaviour in the workplace later faced conflict, hostility and withdrawal of cooperation from perpetrators and colleagues. Such negative consequences punish those speaking out deterring others and reinforcing a culture of silence and tolerance. Where instances of sexual violence go unreported and unpunished for whatever reason the silence maintains the impression that such events are unusual and strengthens the taboo against speaking out (Delker et al., 2020, pp. 14, 15 (in pdf reader)).

Prior to the 1970s the lack of a term to describe sexual harassment as a wrong was a barrier to communicating the experience (Fricker, 2007, p. 151). Several earlier tropes such as the predatory boss (Figure 9) referred to sexual harassment but usually framed it as

humorous and acceptable. Unacceptable behaviours were often positioned as flirting, or the recipient was accused of lacking a sense of humour (Fricker, 2007, pp. 153, 155). Fricker uses the phrases: *hermeneutical injustice* to describe the absence of a word that would enable a group to comprehend a wrong done to them (Fricker, 2007, pp. 147-175); and *hermeneutical marginalisation* involving incorrect stereotypes that create a barrier to particular jobs

(Fricker, 2007, p. 154). Unable to articulate what has happened the victimised are prevented from protesting or preventing a recurrence.

### **2.5.10 Linguistic hi-jacking of VICTIM**

Silencing may also be achieved by devaluing the words used to challenge wrongs. VICTIM did not always carry derogatory meaning and a possible mechanism for this development is provided by Anderson (2020, pp. 2, 3). He suggests that a key contested word may be taken up and used against a less dominant group in a process called linguistic hi-jacking. Here a term is distorted or weakened purposefully to foster active ignorance and resist knowledge in order to “protect systems of domination”. For example the use of the word VICTIM by anti-victimists to counter-accuse those alleging sexual violence<sup>29</sup> of defamation using phrases such as *You are not the victim here – I am!*, foregrounding the accused as the “real victim” as described by Jones et al (2022, p. 6) in relation to the #NotAllMen hashtag. The reversal incorrectly sets up the accusation as equivalent to the violence itself and subverts the meaning to reposition the accused as the “real victim” (Shoaf et al., 2024, p. 1910). The damage done to linguistic and conceptual resources, such as the powerful historic sense of victimisation associated with VICTIM (Anderson, 2020, pp. 15, 16) obscures understanding of the individual’s experiences and prevents them from communicating that to others (Maitra, 2018, p. 345).

Chouliaraki (2024, p. 24) suggests that the sense of outraged victimhood seen in reversals arises from the loss of something the accused felt entitled to, but the reader should ask why the claimant’s:

*...“suffering [is] worthy of a response? Which social position is he speaking from? Which group gains from his claim? Importantly, who loses out?”*  
(Chouliaraki, 2024, p. 24)

Linguistic hi-jacking is discussed again in section 7.3.6. The next section discusses the assignment of responsibility.

### **2.5.11 Depoliticising victimisation**

On a societal level Delker et al (2020, p. 4 (in pdf reader)) proposed that there are two contradictory expectations acting on sexual violence survivors. Firstly, that they will be silent about their experiences, but secondly and in contradiction that they will tell others an empowering and positive transformation story. Individualistic western cultures tend to favour solutions involving personal responsibility (Christie, 1986; Cole, 2007; Sehgel, 2016). This is a view supported by the mental health industry’s interest in *positivity* and *resilience* in therapy as these attributes shift responsibility for the issue onto the victimised (Fisher &

---

<sup>29</sup> This reversal of an accusation is demonstrated in recent claims that the Black Lives Matter (BLM) movement is racist, usually accompanied by counter claims of either *All Lives Matter* or *White Lives Matter* in what is termed Linguistic Hijacking (Anderson, 2020, p. 2). However, the affirmative action in BLM is not racist because non-white populations are subject to systematic disadvantage that white people, the dominant group, are not. This incorrect use of words, whether *racist* or *victim*, by dominant groups to create resistance to the understanding of oppression, effectively becomes linguistic hijacking. Anderson’s main theme was racism, but included victimisation as an adjunct.

Jones, 2024, p. 186). They point to a dominant discourse of recovery represented as a personal responsibility “excusing structural causes of distress” (Fisher & Jones, 2024, p. 186). Much earlier Mahoney (1995, p. 62) suggested why acknowledgement of structural or social causes might be resisted by a culture; because that would reframe individual victimisation as oppression.

Moving on to consider language itself as a deterrent to speaking out in the first place the next section discusses hate speech.

### **2.5.12 Victimization by hate speech**

This section describes how hate speech may intimidate the victimised into silence and I will return to this in section 6.6.

Jane (2014, pp. 531, 532) describes as “e-bile” the gendered online bullying that includes trolling, misogyny, and “recreational nastiness” which disproportionately targets and attempts to silence women especially those offering an opinion. Trolls rarely seek to argue a point but instead use e-bile to raise the “discursive volume” (Jane, 2014, p. 534). She found that both women and men receive hate speech but women perceived threats as greater and more disturbing. She cites evidence from 1999-2012 English language media that predominantly involved journalists and academics speaking out on topical issues. The e-bile contained credible threats of murder, rape or other violence to the woman, her family, or pets, along with evidence that the troll knew her address. As a consequence individual women reduced or changed their online engagement and some withdrew from real world activities, with damaging career implications (Jane, 2014, p. 536). Jane found the majority of research into trolling “downplays, trivialises, or explicitly celebrates e-bile” (2014, p. 540).

Abusive hate speech is a characteristic of populism which needs to construct an enemy against whom supporters can unify and contrast themselves (Pardo, 2024, p. 328). By creating “fear, anger and rejection” of the Other a sense of “ideological right hegemony and belonging in the citizens’ own social strata” (Pardo, 2024, p. 328) is created benefiting the troll’s in-group. This suggests a negative example of Zappavigna’s ambient affiliation discussed in section 2.3.2.

## **2.6 Conclusion**

This chapter has situated the debate on sexual violence and the gender pay gap, described the methodology, mapped the relevant sociological landscape, and provided some background on the VICTIM and SURVIVOR labels. In the rest of this thesis I will build on this research by:

- Applying FPDA (2.3.2) concepts to large-scale deduplicated corpora of text from both sexual violence and gender pay gap texts;
- Using JWT (2.4.4) to explain the persistence of rape myths discussed by FLT (2.4.3);
- Propose a hermeneutical injustice within the pay gap corpora;
- Using corpus analysis to strengthen and quantify Feminist Legal Theory (FLT) claims with evidence from both sexual violence and pay gap corpora.

The literature discussed in this chapter borrows widely from areas including Linguistics, Philosophy, Social Sciences, Psychology, and Criminology to provide the rich theories and insights enabling this study's attempt to fill the niche identified in Section 1.2.

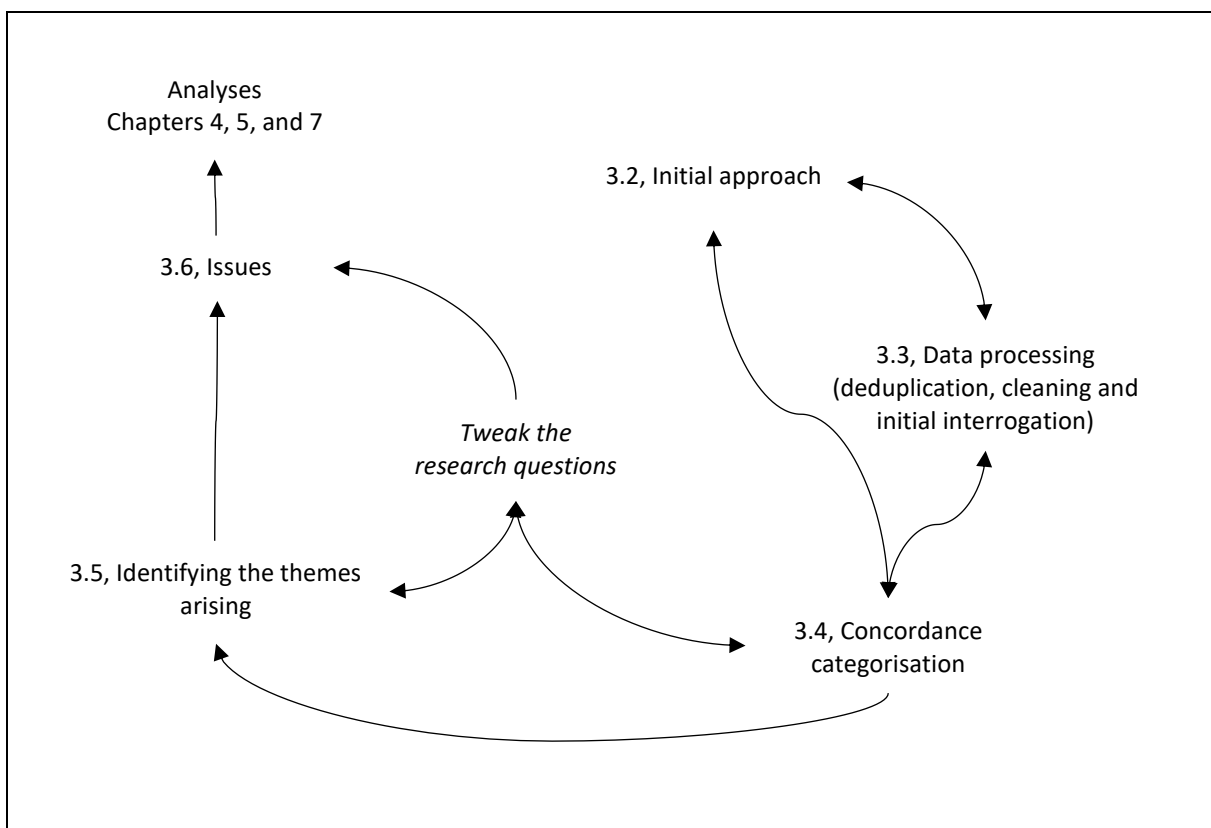
The following chapters explore tweets and news articles to investigate possible links between sexual violence and the gender pay gap language but first the next chapter discusses the methodology.

### 3 Methodology

#### 3.1 Introduction

This chapter’s purpose is to clearly state the thinking and decisions reached at each stage of the study and the thought processes involved. Designing the methodology was an iterative and organic exploration. Along the way the research questions were reworked, methods trialed and tweaked, and themes pursued and abandoned. The complexity of real world data necessitated a flexible approach as accommodating this “messiness” is the price paid for the unexpected findings that emerge and the opportunities they represent (Gillings et al., 2023, p. 43). The description within this chapter simplifies that journey laying out the decisions about topics, tools, and data that lie on the critical path leading into the analyses, illustrated in Figure 10.

Figure 10, Methodology path  
The numbers refer to the section headings of this chapter.



The initial approach providing a way into the subject is outlined in section 3.2. Section 3.3 describes the processing of raw data before sections 3.4 and 3.5 discuss the manual assignment of concordances to categories and themes respectively. Having processed the data, section 3.6 outlines the issues encountered along the way.

Appendix B shows the naming conventions used for the corpora and datasets.

## 3.2 Initial approach

The focus of this study is on the individuals who are victimised by sexual violence or wronged by the pay gap. In the context of sexual violence this generally means those labeled VICTIM and SURVIVOR. There is no equivalent term for those who are wronged by the pay gap and more commonly it is the pay gap phenomena itself that is discussed rather than those disadvantaged by it. Hence, representations are sought beyond those explicitly represented to those who are implied or expected to be involved. This brings into focus the missing or silenced who turn out to be important to this study<sup>30</sup>.

The rest of this section discusses the decisions made in the early stages of this research.

### 3.2.1 The Twitter and Nexis data

This study is based on Twitter and Nexis data, with Nexis news articles providing both a synchronic comparison of more conventional text, and helping to explain the more cryptic tweets on the same topics.

Twitter initially suggested itself because it is the main channel for the #MeToo movement but also because it is a dynamic site where people are “negotiating values, and enacting identities and communities” (Zappavigna, 2017, p. 201). The second element of my research, the pay gap, is also a popular topic on Twitter strengthening the choice of this platform.

The initial intention was to use either a #MeToo or a sexual harassment corpus, expecting that they would be similar. However comparing downloads of *#MeToo* and *sexual harassment*, search results showed surprisingly little overlap. Only 3% of *sexual harassment* tweets contained a *#MeToo* reference, and 2.5% of *#MeToo* tweets containing the phrase *sexual harassment* (Table 2<sup>31</sup>). Furthermore only 1% of @useraccounts appeared in both corpora. This low level of overlap suggested the possibility of different patterns of meaning constructed within each of these two corpora. *#MeToo* tweets showed a number of distinctive features such as a higher proportion of English as second language and specific national interests. These included: Bollywood celebrities (India); K-Pop celebrities (South Korea); and KuToo<sup>32</sup> activism (Japan). Consequently both search terms are analysed going forward.

---

<sup>30</sup> Discussed in section 3.5.2.

<sup>31</sup> Initially using monthly corpora, but updated with \_1Yr figures shown here.

<sup>32</sup> KuToo is activism against the requirement for women (but not men) to wear high heeled shoes in some workplaces.

Search term	TwitterSH_1Yr (42,592 tweets)	TwitterMT_1Yr (47384 tweets)
SEXUAL HARASSMENT	(occurs in 100% of tweets)	1189 (710.67 pmw) occurs in 2.5% of tweets
#MeToo	1,280 (817.51 pmw) occurs in 3% of tweets	(occurs in 100% of tweets)

Table 2, *Overlap in occurrences of SEXUAL HARASSMENT and #MeToo TwitterSH\_1Yr and TwitterMT\_1Yr corpora.*

Those wronged by the pay gap are generally referred to as *workers, employees, or women*, which does not distinguish them from those who are not wronged. However, most commonly they are not referred to at all. Consequently, the single search term *pay gap* was used with the hope that analysis would uncover the missing representations.

In order to provide a comparison with the tweets<sup>33</sup>, a source of more conventional discussion on the same subjects was needed. This was provided by the Nexis database English Language news articles from around the world. These longer-form electronic texts are written by professional journalists in the editorial style of their newspaper.

The full 24 month corpora used in the analysis chapters are based on the search terms *sexual harassment, #MeToo*, and *pay gap*.

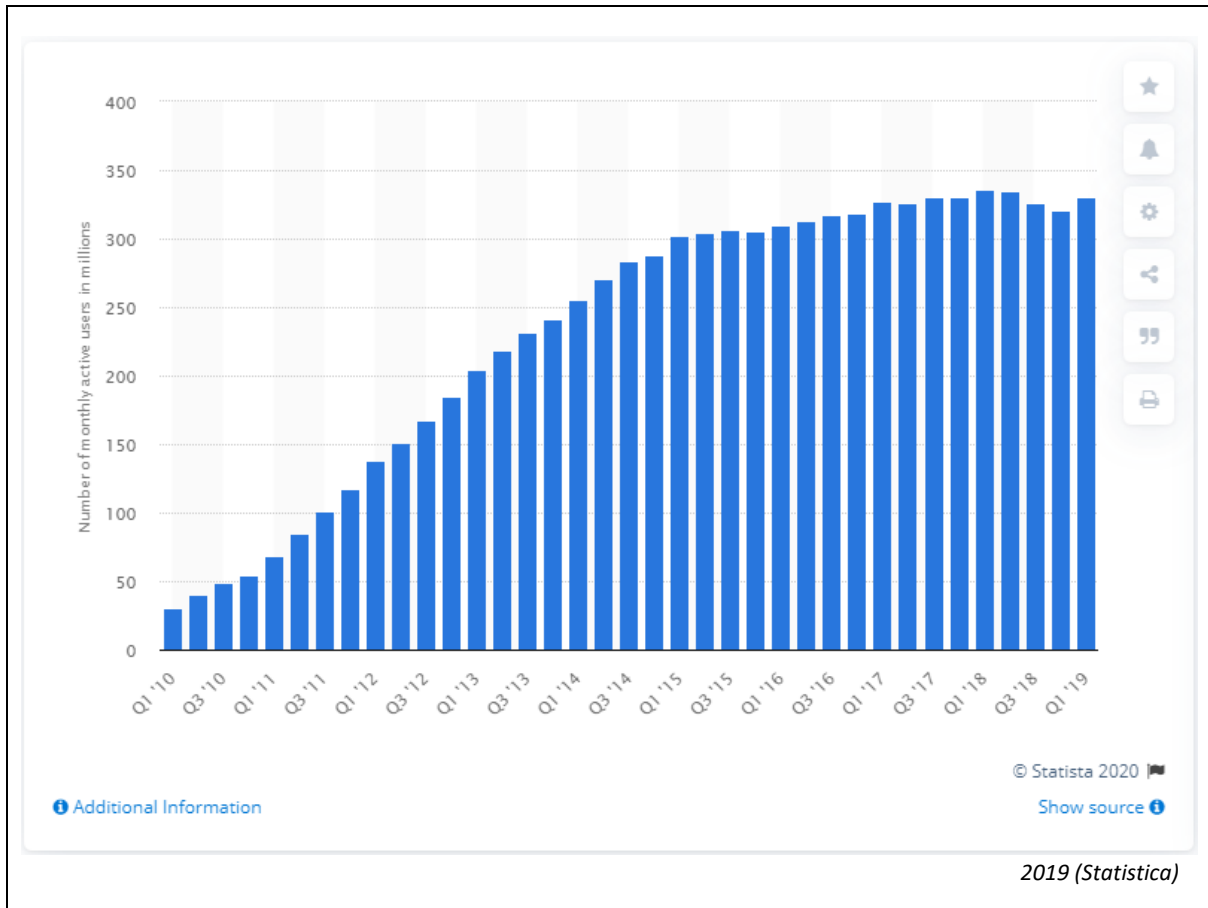
### 3.2.1.1 Twitter

Twitter was launched in 2006 and in the ten years 2009 to 2019 the number of active users grew from 30 million to 330 million (Figure 11) and was predicted to grow by a further 0.3% in 2020. An additional 500 million people were viewing Twitter each month without logging into an account.

Much research has been done on the influence of Twitter on its users in terms of purchasing behaviour and political attitudes. Twitter's own research (@katieaka, 2016) indicates that user behaviour is affected by tweets, with close to 40% of users attributing a purchase directly to tweets posted by influencers. Influencers are users who deliberately seek to affect other users' behaviours although the tools they use are the same as those available to ordinary users. The Digital Marketing Institute (2020) found that recommendations by influencers almost doubled users' intention to buy a product. This has implications beyond purchasing to political engagement, finance, and the shaping of public opinion more generally. For example, there is evidence that active sharing of content on Twitter enhances users' sense of personal agency, which in turn boosts their likelihood of political participation (Halpern et al., 2017). Twitter has also been seen to influence behaviour driving financial markets, for instance when the White House Twitter account was hacked, with tweets posted saying that the White House was under attack the Dow Jones dropped 145 points in a few minutes (Petter, 2019). These examples suggest that the language used on Twitter has power to change things beyond the webpage.

<sup>33</sup> Posts on Twitter are called tweets.

Figure 11, Active Twitter users worldwide, 1<sup>st</sup> quarter 2010 to 1<sup>st</sup> quarter 2019



Halpern, Valenzuela and Katz (2017) and Clement (2019a) found that activity on Twitter was oriented around users sharing information. While Twitter was likened to news broadcasting (Gorbach, 2016) in contrast to Facebook which was described as akin to a tea party where users are encouraged to engage with all others as friends. The purpose to which users put social media platforms affects the type of content they generate and that which they seek to access. According to a study by Business of Apps (2020) for the year ending June 2019 59% of account holders use Twitter to access news. The extent of Twitter’s news orientation is seen in Figure 12 where at the beginning of February 2019 Twitter users engaged in news sharing activities 50% more than those on Facebook, the next most news-focused platform. Consequently, it is not surprising that 25% of verified Twitter accounts are held by journalists (Business\_of\_Apps, 2020). Many other Twitter accounts are not easily identifiable as particular individuals or organisations and this anonymity and consequent unaccountability may colour their contributions.

Twitter is often associated with political activism, causes, and campaigning, which at their extremes can deteriorate into inappropriate and aggressive behaviour (Rose & McGrory, 2020). During 2019 31% of American users engaged in discussions about national politics via Twitter (Business\_of\_Apps, 2020) making the platform a useful window into opinions on both local and global topics for citizens, politicians, professional communicators and world leaders (Clement, 2019a).

Figure 12, Comparing social media activities, February 2019

**Social media activities on select social networks by social media users in the United States in February 2019**

	Instagram	Facebook	Snapchat	Pinterest	Twitter
News	18%	38%	17%	9%	56%
Viewing photos	77%	65%	64%	59%	42%
Watching videos	51%	46%	50%	21%	32%
Sharing content with everyone	45%	57%	46%	21%	32%
Networking	23%	33%	21%	10%	26%
Sharing content one to one	31%	43%	45%	12%	20%
Finding/shopping for products	11%	15%	5%	47%	7%
Promoting my business	9%	7%	6%	5%	7%

Source: Statista

(Statistica)

According to Clement (2019b) the average duration of a Twitter session was 3.53 minutes suggesting Twitter users do not spend long deliberating over their wording. Furthermore, the home page prompt “What’s happening?” encourages a spontaneous response that may be closer to everyday speech. Alongside this the maximum tweet length of 280 characters ensures that tweets are economical with words, contributing to the different characteristics of Twitter’s language from mainstream news articles.

**3.2.1.2 Nexis**

Nexis (n.d.-c) is an on-line searchable archive of news and business article that provides access to 40,000 sources. It allows articles returned for a search term and selected period to be downloaded as MS Word documents that can then be manipulated to create corpora.

Within the Nexis text downloads there are titles from around the globe ranging from well known national newspapers with high circulation figures down to titles focused on a single city, industry, or university. Subject matter may be general interest or specialist.

My decision to use Nexis was based on both the free access granted via Lancaster University, and on my experience with it earlier in my studies meaning it represented no learning curve. The use of Nexis to harvest news articles is described in section 3.3.3.

### 3.2.2 The raw data

The present study's data was gathered from a new Twitter account with a blank baseline search algorithm on a PC that had not accessed Twitter previously. Most users would normally have a string of "Follows" and a history that influences the posts selected for their feed. Although free to users, Twitter like all social media is a commercial concern and needs user engagement to drive traffic and attract and keep advertising revenue. Milli et al (2025, pp. 2, 6) found that Twitter's algorithms amplify existing biases toward attention-grabbing tweets that are more partisan, outgroup hostile, and extreme, than expected from user preferences. Finkel et al (2020, p. 534) allude to likely user responses to "the contagious power of content that elicits sectarian fear or indignation". Tweets and news articles were collected monthly for 24 months starting with April 2019. Each month's corpora were prefixed with digits 01 to 24 to allow for chronological analysis if needed.

The purpose here was not to create a general representative corpus but to "afford a rich investigation of meaning-making" on the topics of interest (Zappavigna, 2012a, p. 8 (in pdf reader)) based on the search terms *sexual harassment*, *#MeToo*, and *pay gap*. The #MeToo data contained a much higher percentage of tweets comprised exclusively of non-English language text, emojis, URLs, hashtags, and @usernames. Furthermore, the English language text in #MeToo tweets also indicated input from a wider range of nationalities than the sexual harassment corpora. The degree of difference was reduced after data cleaning but still showed that the sexual harassment and #MeToo corpora were qualitatively different. Hence, a parallel #MeToo download was taken from Nexis.

Nexis "All English Language News" data provides a good comparison to Twitter as it can be discussing the same situations but using more formal language with more in-depth coverage providing insight into topics referred to only cryptically in tweets.

There are practical differences between the downloaded Nexis articles and the tweets making separate methods of data cleaning necessary. For example the 01-12 monthly corpora contained an average of 29 words for the sexual harassment tweets, 32 for pay gap tweets, and 26 words for #MeToo tweets (Table 3). The Nexis news articles also showed lengthier coverage of the pay gap articles at an average of 1,097 words compared with the averaged 964 and 899 words for sexual harassment and #MeToo respectively.

	SEXUAL HARASSMENT	#MeToo	PAY GAP
Twitter	29	26	32
Nexis	964	899	1,097

Table 3, Average number of words in tweets/articles, based on the \_1Yr corpora

The content of the two media sources is also very different, Nexis texts being formal one-to-many professional writing by a named journalist and newspaper who could be held accountable, while tweets are informal, reactive, and often not attributable to an individual. A random one hundred tweets were checked for potential @useraccount identification (first name/initial plus family name, an e-mail address, or company or organisation name). This

revealed considerable variation between the pay gap at 88% identifiable and sexual harassment at 41% identifiable (Table 4). Surprisingly #MeToo tweets (78%) were almost twice as likely as the sexual harassment equivalents to be identifiable.

Tweets originating from @useraccounts that are not clearly linked to traceable individuals carry implications for the accountability of truth and accuracy.

	SEXUAL HARASSMENT	#METOO	PAY GAP
Twitter	41%	78%	88%

Table 4, Tweets in a randomised 100 that had a potentially identifiable @useraccount

From this point onwards when sexual harassment and #MeToo data are discussed together they are referred to as sexual violence corpora or datasets.

**3.2.3 The reference corpora**

The British National Corpus (BNC) (Consortium, 2022) was used as a reference corpora in preference to the more recent BNC 2014 because at the time of my research, the written component of the BNC 2014 was not available to me. The BNC is an English Language corpus of 100 million words taken from late 20<sup>th</sup> century written (90%) and spoken (10%) sources. It is a general corpus with samples taken from a wide range of non-internet genres, with the most common text type being written books and periodicals. The BNC was used when a comparison with general usage language was required.

In contrast, the enTenTen corpus (SketchEngine, n.d.) is made up entirely of English language text harvested from the internet. The enTenTen21's (52 billion words) text was harvested in 2021 and the largest genres are culture/entertainment, travel, and religion. Initially this study used the enTenTen15<sup>34</sup> (13 billion words) (Chapter 3) harvested in 2015, but shifted to the enTenTen21 when that became available (Chapters 4-7). The enTenTen corpora were used when a comparison with more general internet language was needed.

**3.2.4 Who are represented in the sexual violence and pay gap corpora?**

The people represented in the sexual violence and pay gap corpora are predominantly women (Table 5) although their relationship to discrimination cannot be identified here.

Together the lemmas WOMAN, GIRL, GENDER, and FEMALE, (labelled the Woman Group in Table 5) are used between 1.7 and 4.5 times more frequently than the lemmas MAN, MALE, and GUY, (the Man Group in Table 5) and between 13 and 19 times more frequently than the lemmas ETHNIC, BLACK, ASIAN, LATINO, MINORITY, and RACIAL (the Ethnicity Group in Table 5). Individuals may be members of more than one demographic group but in tweets and news media they are often identified as belonging to only one group. For instance Example 216 refers to "... not discriminating against anybody - black, gay, religious, female, etc." and Example 24 which offers "Accuse the blacks, play VICTIM and get funded" which suggests that the single attribute may have been identified as of greater significance than a random selection.

---

<sup>34</sup> Genre information is not available for the enTenTen15.

	NexisSH_1Yr	TwitterSH_1Yr	NexisMT_1Yr	TwitterMT_1Yr	NexisPG_1Yr	TwitterPG_1Yr
<i>Woman Group pmw</i>	10,394.22	5870.23	3,631.73	6,055.91	11,189.83	22,872.81
<i>Man Group pmw</i>	3,330.36	2,435.23	1,602.09	3,627.62	3,587.27	5,061.7
<i><u>Woman words per M</u></i>	3.12	2.4	2.27	1.67	3.12	4.52
<i><u>Man pmw</u></i>						
<i>Ethnicity Group pmw</i>	660.55	309.60	648.148	444.65	689.75	1,185.49
<i><u>Woman words per M</u></i>	15.74	18.96	5.6	13.62	16.22	19.29
<i><u>Ethnicity pmw</u></i>						
<i>Disabled pmw</i>	105.11	4.61	4.6	14.32	108.66	66.60
<i><u>Woman pmw</u></i>	98.89	1.273.37	789.5	422.90	97.62	346.24
<i><u>Disabled pmw</u></i>						

Table 5, The most common targets of sexual violence or pay disparity pmw in the \_1Yr corpora

That VICTIM and SURVIVOR are more commonly associated with female rather than male nouns (Messamore & Paxton, 2021, p. 16) creates a female prosody around the terms that may challenge the stereotypical male identity, creating a barrier to men speaking out. This would make interesting research but is beyond the scope of this study.

Across the pay gap and sexual violence corpora there are some common features: sex is the most frequent identifier of individuals; ethnicity is less frequent but still significant; and disability relatively rare. Interestingly the Twitter pay gap corpora has much more frequent references to Woman Group lemmas than the other corpora in Table 5.

The news media corpora from Nexis focuses on a few celebrity cases, such as Carrie Gracie's (BBC, 2018a) equal pay case against the BBC, and the inquiry into Christine Blasey-Ford's (BBC, 2018b) testimony of sexual harassment against Brett Kavanaugh in the US. The parallel Twitter corpora feature the same celebrity cases, but also large numbers of tweets involving other less well-known individuals not reported in Nexis news articles, and also anonymous cases. The Nexis sexual violence corpora most commonly use the term VICTIM rather than SURVIVOR reflecting the legal use of the word in court case reporting.

The distinction between VICTIM and SURVIVOR and who may and who may not be referred to as such, is often positioned as a very distinct and deliberate choice. An early finding was that Twitter sexual violence texts contain interesting reversals that position the original accuser as bogus and having deliberately fabricated their case, repositioning the original accused as the *real victim*. Reversals are less frequent in the Nexis corpora, however NexisPG\_1Yr does

contain claims that the gender pay gap is a fabrication claiming disadvantage where none exists. Reversals never refer to an accused as the *real survivor*.

Cases of pay disparity within Nexis articles are more likely to refer to broad demographic groups such as *employees, women, Hispanics, etc.*, rather than individuals. As a statistical term *pay gap* is discussed in terms of populations or demographic groups, however when the topic is *equal pay* references are more associated with individuals or groups. The terms used never suggest injustice in the way that SURVIVOR or VICTIM do. Another distinguishing feature of the pay gap texts is that the issue is generally current and on-going.

#### **3.2.4.1 Alternative terms**

To ensure that other terms for victimised individuals were not overlooked several related words were trialed on the \_03\_2019\_06Jun family of corpora. Some of the instances picked up were within a phrase that was not relevant to sexual violence or the pay gap topics and these were subtracted from the numbers (Table 6). Such instances included lists of unrelated article titles featuring elsewhere in the source publication. Additionally, within the NPG\_03\_2019\_06Jun corpus of 80 instances of VICTIM only 19 applied to pay gap situations and the other 61 were actually referring to sexual violence.

*Worker* and *employee* are both emotionally neutral terms that are associated with legal contracts and workplace rights. The other lemmas trialed (Table 6) carry meanings involving justice, the degree of harm, and agency. NSH\_03\_2019\_06Jun returned significantly higher numbers of *worker* (768 pmw) and *employee* (1,807 pmw) compared with TSH\_03\_2019\_06Jun (562 and 798 pmw respectively), or TMT\_03\_2019\_06Jun (163 and 217 pmw respectively). This suggests greater focus on workplace issues which may reflect the greater visibility and potential evidence within a public environment. The same pattern is repeated for NPG\_03\_2019\_06Jun which has higher frequencies of *worker* (1,417 pmw) and *employee* (1,471 pmw) than TPG\_03\_2019\_06Jun (1,291 and 773 pmw respectively). For the other lemmas (Table 6) Twitter corpora return higher frequencies than their Nexis counterparts except for the word *prey* which occurs at a lower frequency in the TSH\_03\_2019\_06Jun (18 pmw) than in the NSH\_03\_2019\_06Jun (38 pmw) corpus.

Lemma	NSH_03_2019_06Jun pmw	NMT_03_2019_06Jun pmw	TSH_03_2019_06Jun pmw	TMT_03_2019_06Jun pmw	NPG_03_2019_06Jun pmw	TPG_03_2019_06Jun pmw
VICTIM	1,332	631.9	1,659	1,195	19	104
SURVIVOR	206	210.63	294	931	4	7
INNOCENT	50	84.25	140	316	0	10
CASUALTY	0	0	0	13	0	1
PREY	38	16.85	18	63	1	1
TARGET	20	126.38	170	178	0	4
WORKER	768	151.66	562	163	1,417	1,291
EMPLOYEE	1,807	294.89	798	217	1,471	773

Table 6, Potential labels for the victimised for the \_03\_2019\_06Jun corpora.

This lack of visibility of the disadvantaged initially presented a challenge so the pronoun I which was the highest frequency pronoun in most of the corpora was chosen for further exploration to try to identify some personal statements. The concordances for I was within the \_03\_2019\_06Jun corpora were downloaded and manually encoded for the roles represented: *Victim*; *Accused*; *Commentator*, or the *Accused as victim* (a reversal) (Table 7). Within the pay gap corpora most frequently it was the issue itself being discussed rather than specific instances, with companies, industry sectors and countries being the most frequently named entities. Within TwitterPG the collocates of *salary* include three specific employers: BBC (9), NBA (4), and WNBA (4) which across the corpus (collocates plus non-collocates) occur 500 times. In contrast, the three generic employer collocates: employers (4), board (3), and companies (4) across the corpus (collocates plus non-collocates) occur 1,397 times. Those responsible for paying salaries are more often represented as distanced and generalised.

Social actor role	NSH_03_2019_06Jun	TSH_03_2019_06Jun	TMT_03_2019_06Jun	NPG_03_2019_06Jun	TPG_03_2019_06Jun
Total I WAS...	46	43	69	28	22
Victim	76%	58%	70%	89%	32%
Commentator	15%	28%	14%	7%	9%
Accused	7%	2%	6%	0%	27%
Accused as Victim	2%	0%	0%	0%	0%
Other	0%	12%	14%	4%	27%

Table 7, Breakdown of social roles within concordances for I WAS... for the \_03\_2019\_06Jun corpora.

Within all corpora *Victim* was the most common role represented and *Accused as Victim* the least common. Generally *Commentator* was the second most common with *Accused* the third. The exception being TPG\_03\_2019\_06Jun where this order was reversed. Interestingly the high frequency of *Victim* voices within the NPG\_03\_2019\_06Jun corpus (89%) was driven by articles quoting victims describing the co-occurrence of sexual violence alongside pay gap issues. This indicates that the pay gap is seldom discussed with reference to those victimised by it highlighting the need for a different approach. The subsequent work identifying themes continues in section 3.4.2.1.

**3.2.4.2 Quid pro quo**

Discussions about the pay gap and sexual violence sometimes occur in association with quid pro quo cases (something for something) where sexual violence is portrayed as fair trade involving a fully aware and agentic victim. Contrasting perspectives on the same situation may be raised as the following examples show: a) represents Weinstein’s victims as having power in the relationship and willingly trading sexual favours to further their careers; and b) illustrates the opposite situation where the powerless victim is threatened with damage to their career if they do not provide sexual favours.

- a) Exactly. Weinstein is just a freaky old man who used his name & status & these women gladly participated in the “casting couch” game. No grown woman who has been genuinely raped by a man writes emails like “Miss ya big guy!” These old groupies jumped on the #metoo bus 4 attention! (TwitterMT\_2019\_02Feb)
- b) On the other hand, quid pro quo sexual harassment is the grimy, Weinstein-esque routine of “perform a sexual favor for me, date me, etc. if you want this job or want to keep your job, or you want your mom to keep her job.” (TwitterSH\_2019\_07Jul)

This presents a risk of counter-accusation to women speaking out about sexual violence within a work situation. Example b) suggests an explicit verbal contract stating terms,

however it is more likely to be implicit and open to interpretation, so may mean different things to the two parties.

The writer of a) mistakenly believes that all victims of sexual violence behave in the same way, and also overlooks the likelihood of fawning (Clayton, 2025, pp. 21, 24), in which a victim may need to maintain a relationship with a perpetrator particularly where a power differential exists.

Quid pro quo situations are a well-established phenomenon historically combining transactional sex and usually massive power differentials. Shakespeare's (2001, p. 235) Henry the VI part 3 has King Edward deliver the line "Now tell me, madam, do you love your children?" which seems uncomfortably close to example b) above. In this situation Chouliaraki (2024, p. 36) would advise readers to question, "Which positions of power do these selves speak from? What kinds of benefits do their claims accrue to them and the communities they bring together?".

### **3.2.5 A feminist perspective**

In theory, the Equality Act 2010 (2010) provides protection for all demographics against a broad range of discriminatory practices including sexual violence and pay disparity. However, the evidence of this study suggests that this protection is not fail safe with many groups still negatively affected. As the targets of these types of discrimination are predominantly female this study leans toward a gendered perspective. Since the 1970s much overt discrimination has been eliminated or reduced and equality is now generally viewed as a good thing. Tranchese (2023, pp. 289, 321) found no strong association between overt victim-blaming and rape discourse, and that which she did find was more often recognised and publicly condemned. However, #MeToo has highlighted that less visible cultural and political beliefs remain that result in the remaining inequalities being seen as normal and natural (Tranchese, 2023, p. 338).

While I have taken care not to bias the data and analysis, it is inevitable that some of my own values as a female, white, mature, middle class, UK citizen, will be brought to the process. My interpretations may differ from those of another researcher reflecting our individual awareness of self and also awareness that the discourse reflects the "commonsense" of the prevailing ideology (Fairclough, 2001, p. 64).

### **3.3 Data processing**

This section describes the mechanical processing of data and the issues that arose. Tweets were gathered on four occasions each calendar month and where this did not generate enough processed tweets a fifth collection was made ensuring that all monthly corpora were approximately 100,000 words. The SEXUAL HARASSMENT and #METOO search terms always generated large volumes of data that were all from a single day. In contrast smaller numbers were always returned for the PAY GAP search term and these would represent postings over a number of days. In one month during the first UK lockdown of the Covid19 pandemic the final pay gap corpora contained all the unique tweets posted that month. News articles were gathered monthly for the previous calendar month with 300 articles downloaded to ensure a final monthly corpus of 100,000 words. At the end of the twelve-month period each set of 12 monthly corpora were combined and deduplicated to create six 12-month corpora of

approximately 1.2 million words. Monthly processing of the data was a necessary step to ensure that the larger 12-month corpora did not overload the software. This also provided a useful preview of the data for pilot analyses.

A range of different tools were needed to facilitate the various stages of data processing: harvesting raw data; cleaning the data; deduplication; and interrogation of the resulting corpora. Due to the different nature of electronic news articles and tweets different tools were also needed for Nexis and Twitter data. As the intention of this study was to analyse text rather than explore technological solutions, tools were selected on the pragmatic basis of availability and a short learning curve.

The following sections discuss the reasoning for choosing the data sources and tools. Processing of the Twitter data is described in section 3.3.1 and 3.3.2, and Nexis data in sections 3.3.3 and 3.3.4.

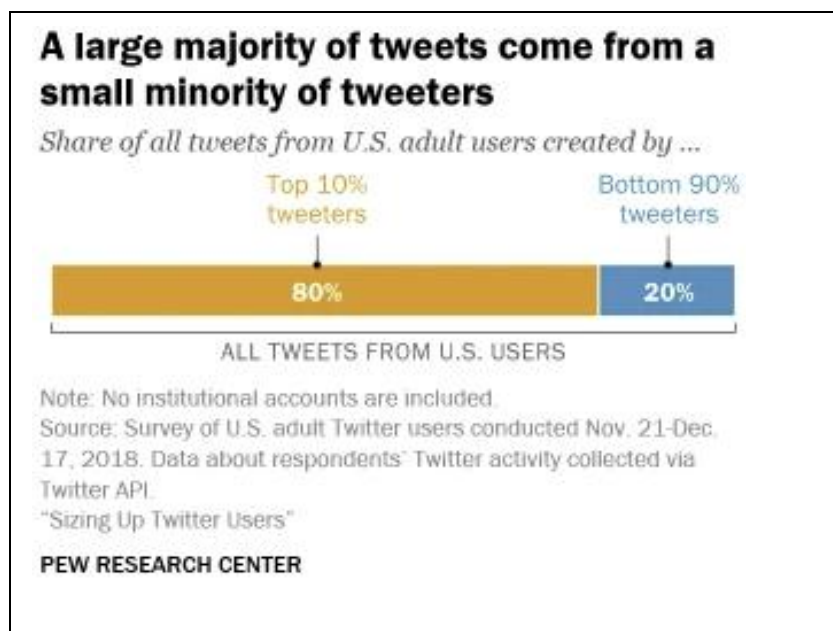
### 3.3.1 Harvesting tweets (TAGS)

The Twitter Archiving Google Spreadsheet (TAGS v6.1.9) (Hawksey, n.d.) accesses the Twitter Search API to harvest tweets posted during the last seven days, capturing tweets in reverse chronological order from the date selected. The search terms used were PAY GAP, SEXUAL HARASSMENT and #METOO, with Retweets excluded, resulting in a collection of relevant indexed tweets (Twitter, n.d.). In a few seconds 5000 tweets were collected for download in .XLSX format. The predominant countries of origin of the tweets varied with the time of day so tweets were routinely captured between 06.00 and 10.00 hours to aid consistency.

One of the features of Twitter is that the heavy use of retweeting results in a few individuals having a disproportionately strong influence. The Pew Research Center<sup>35</sup> found that, based on their survey of 2,791 US adult Twitter users, 80% of tweets were created by only 10% of users (Wojcik & Hughes, 2019) (Figure 13). This was reflected in the raw data of this study which contained high volumes of identical retweets and very close

Figure 13, 80% of tweets authored by 10% of users

(Wojcik & Hughes, 2019)



<sup>35</sup> The Pew Research Center is a non-partisan organisation that undertakes and publishes data driven social research based on public opinion polling (Pew Research Center, n.d.).

matches perhaps with just an additional word or hashtag added.

The TAGS function for removing duplicates was unable to remove the close but not identical tweets with only minor variations such as a single different URL, @username, or emoji.

Deduplication of Twitter @useraccounts was also carried out so that each @useraccount contributed no more than a single tweet to each monthly corpus preventing a few prolific users dominating the discourse. Further deduplication was carried out during aggregation to remove any duplication between monthly data files, and this is described in section (3.3.6).

### **3.3.2 The deduplication and cleaning process for tweets (Excel)**

Raw data downloaded from Twitter is messy with much duplication, as one of Zappavigna's examples put it "User 2: ... akin to thousands of unedited Anne Frank diaries, a huge pile of telegraph cables, many out of context or relation" (2012b, pp. Chapter 9, paragraph 3 (in pdf reader)). Hence, both deduplication and data cleaning are a necessity. Without deduplication the corpora would have been dominated by a few high profile accounts with millions of retweeting followers. As a consequence of deduplication popular words and phrases that may have gone viral on the Twitter platform are much less frequent within these corpora. Instead, the more modest peaks of frequency captured here result from a number of different writers, many of whom will not be journalists or other high profile individuals. This study can say nothing relating to the frequency of features as they would have been viewed in Twitter.com or news websites at the time.

Because tweets are limited to 280 words and many were much smaller (Table 3) they lend themselves to viewing in an Excel spreadsheet. Excel formulae provide a flexible way to carry out further deduplication and data cleaning. Tweets containing less than three words of English were removed as it was difficult to make meaning from these. Most tweets larger than 400 characters were found to be either entirely or predominantly lists of hashtags. To avoid distorting the statistics, tweets consisting entirely of #hashtags, @usernames, URLs or emojis, which Schäfer and Bildhauer (2013, pp. 71, 72) calls "noisy data" were also excluded. However, repeating orthographic patterns involving punctuation were retained. This cleaning operation and subsequent randomisation was achieved with a series of Excel formulae (Appendix C). Excel was also used during coding and analysis to enable annotation, sorting, and filtering.

While Nexis allows selection to be restricted to English Language articles all Twitter downloads contained some non-English language tweets, and this was particularly common in the #MeToo downloads. As this study intended to study only English Language tweets it was necessary to remove as many non-English language postings as possible. This was done by creating Excel formula to highlight rows containing words from a list of common non-English language words and characters used in the tweets (Appendix D). After each iteration using the existing list of non-English language words the remaining tweets were searched again for a new commonly occurring word to add to the list. When this list was stable the formulae was run on all corpora. A second formula was created in a similar way containing a list of diacritic characters and a third with non-alphabets such as South Korean ideograms.

It was not possible to deduplicate the Twitter files in the same way as Nexis files described in section 3.3.4, because the large number of small files<sup>36</sup> and high rates of closely similar texts caused WordSmith to crash. Instead within Excel the first 150 characters of each tweet were isolated and subject to Excel’s Conditional Formatting feature then Find Duplicate Values was used to identify and then remove duplicates.

After cleaning and deduplication the next stage was to insert the marker “Startoftweet\_yyyymmdd” before each tweet and then save the spreadsheet as a text file which could be uploaded to SketchEngine to create a corpus. Any thinning of tweets needed to reduce the size to 100,000 words for the monthly corpora was done from the bottom of the randomised text file.

**3.3.3 Harvesting news articles (Nexis)**

News items were harvested from Nexis using Nexis Power Search (n.d.-a) using the search terms PAY GAP, SEXUAL HARASSMENT, and #METOO in conjunction with the Publication Language set to “English”, and Publication type set to “Newspapers”. The date range was set to the previous calendar month. Results were sorted in “Relevance” order and the first 300 downloaded. This number allowed for subsequent loss during deduplication. Nexis articles required comparatively little data cleaning.

Nexis includes all instances of the articles in downloads, grouping them by similarity. There are three options for grouping: Similarity Off; High Similarity; and Moderate Similarity (Nexis, n.d.-b). When Similarity is off all similar stories can be included within a single group, High Similarity excludes very similar stories, and Moderate Similarity excludes moderately similar stories. Table 8 shows how the results of the different Similarity settings for the search results in the month of March 2020.

	Similarity Off	High Similarity	Moderate Similarity
Pay gap March 2020	2,801	2,747	2,738
% decrease from Similarity Off		1.9% (54)	2.2% (63)
Sexual harassment March 2020	6,688	6,664	6,653
% decrease from Similarity Off		0.4% (24)	0.5% (35)

*Table 8, Effects of Nexis similarity settings March 2020 data.*

To achieve the greatest degree of deduplication a setting of Moderate similarity was used. While the difference was small the Moderate setting identified nine more near duplicates than the High Similarity setting for the pay gap corpora and eleven more for the sexual harassment corpora. After this exercise considerable duplication remained.

A difficulty with commercial news stories is the republishing of the same story in later editions of a newspaper with very minor differences that pass through the Nexis grouping threshold. These may have been syndicated to other titles and republished with only a reworded introduction or conclusion (Alonso et al., 2013). To identify this repetition a further deduplication of the Nexis files was carried out in WordSmith (section 3.3.4).

<sup>36</sup> The #METOO\_1Yr corpora contained 53,000 files.

### 3.3.4 Deduplicating news articles (WordSmith)

Prior to deduplication the MS Word download of articles was split into individual files using the WordSmith Splitter using the “Copyright yyyy”<sup>37</sup> metatext as the “start/end of text separator”. Single article files that differ significantly but are clearly closely related, for instance those containing the same spelling mistakes could then be identified using the WordSmith Duplicate Contents function.

The pilot phase of this process was by trial and error to identify the optimum settings for rigor and practicality. To establish the settings to be used Duplicate Contents was run on the raw data for the NexisPG\_1Yr corpus using a range of settings. Significantly, running time increased with the number of differences between the texts (Table 9). A pragmatic balance was needed between finding the greatest number of duplicates and a practical run-time. This was found to be at a “max length diff %” of 50% and “max words diff %” of 50% (highlighted yellow in Table 9).

Three hundred and two near duplicate texts were found at the most rigorous settings with a total dataset of 1,184 articles. The third column in Table 9 shows the near duplicates found at each run of Duplicate Contents and in the right hand column as a percentage of the 302 found at the most rigorous setting. This showed that the “max words diff %” was the main driver of the number of near duplicates found. While it would have been ideal to have adopted the most rigorous settings reducing the “max words diff %” to 50% was at a cost of missing only 3% of near duplicates but crucially it saved many hours of processing time and occasional WordSmith crashes.

Max length diff %	Max words diff %	Near duplicate texts found	% of words from smaller text present in larger text	% of total (302) near Duplicate texts found
2	2	44	2%	15%
50	2	44	1%	15%
2	50	250	49%	83%
25	25	250	20%	83%
50	25	252	17%	83%
90	25	253	13%	84%
25	50	280	40%	93%
50	50	294	33%	97%
90	50	302	26%	100%

Table 9, Effect of manipulating WordSmith length and difference NexisPG\_1Yr raw data download.

After deduplication the WordSmith Joiner function was run to aggregate the single article files back into a Word document that could be used to create a corpus.

### 3.3.5 Final data cleaning (MS Word)

Even after the processing described in sections 3.3.1 to 3.3.4, the Nexis and Twitter raw data files contained a variety of issues that needed to be addressed including removal of

<sup>37</sup> The current year was used for each file.

unwanted items. For example, within every month's Nexis articles there were a few very long stock reports with only minor and incidental reference to the search terms, and as the bulk of the text was irrelevant and specialist this would have skewed the analysis. These articles were very difficult to filter out by any inbuilt facility within Nexis or WordSmith and the fastest solution turned out to be using the MS Word search facility to identify and then manually delete these.

The last step in preparing texts was to use MS Word to provide a word count as a rough guide to corpus size. The finalised text files were then uploaded into SketchEngine to generate corpora. Along the journey of this study corpora were produced for months, 12 month aggregations, and 24 month aggregation, with the timespans indicated in the naming conventions, listed in Appendix B.

### **3.3.6 Aggregating the monthly data**

Aggregating each set of 12 monthly corpora into a single file raised the possibility of finding a maximum of 12 duplicates. To prevent this each set of 12 files was deduplicated using WordSmith. The general process will be described first, before outlining the differences required for the Twitter files.

Nexis files were split on the "start/end of text separator" as *Copyright yyyy* which ended all Nexis stories. As each run of the Splitter started numbering from zero each of the 12 monthly files was split into a separate folder then the files in all but one folder were renamed using a batch file that allocated each story a unique file name. When this was done all files were saved to a single folder on which the Duplicate Contents utility could be run. The deduplicated contents within the single folder were then aggregated using the WordSmith Joiner utility to create a single 12-month file.

The process for aggregating the 12 monthly Twitter corpora was similar to that used for Nexis with the following exceptions. Having initially deduplicated the Twitter monthly files in Excel the aggregated twelve-month corpora could contain no more than twelve duplicates and this enabled WordSmith utilities to run without crashing. For the Twitter files the WordSmith "start/end of text separator" was set to "Startoftweet\_".

### **3.3.7 Early view of the data**

From the end of the first month of data gathering, corpora were produced to provide a view of the data allowing exploration of possible themes. Early lines of investigation included:

- pronouns and names by accusers, the accused, and commentators;
- use of modals;
- Twitter @useraccount<sup>38</sup> descriptions (who is tweeting); and
- #hash tags.

---

<sup>38</sup> Within Twitter the prefix @ identifies a user account, # a searchable topic. Emojis are characters that signify an emotion.

However, those were discounted because key word searches indicated a heavy bias towards US politics<sup>39</sup> which would make an interesting topic for another study. While modals were a possibility they suggested less scope than the final choice. In comparison, representations associated with pronouns gave a more direct route to the individuals who were affected by sexual harassment and the pay gap.

Within the monthly corpus TSH03\_2019\_06Jun the WordSketch for the pronoun *I* showed that the *I was* structure had the highest LogDice<sup>40</sup> (11.7). Table 10 shows the results for all of the \_2019\_06Jun corpora. The *I WAS* concordances were then manually coded for the social role of the voice.

Social actor roles	NSH03_2019_06_Jun	TSH03_2019_06_Jun	NPG03_2019_06_Jun	TPG03_2019_06_Jun	TMT03_2019_06_Jun
Total frequency I WAS...	46	43	28	22	69
Victim	76%	58%	89%	32%	70%
Accused	7%	2%	0%	27%	6%
Commentator	15%	28%	7%	9%	14%
Accused as victim	2%	0%	0%	0%	0%
Other	0%	12%	4%	27%	14%

Table 10, *I WAS* pilot concordances  
Categorised by social role, based on 03\_2019\_06Jun data  
Note that at this point there was a problem with the Nexis #MeToo data so this was omitted.

Seeing that the *victim* voice was the most frequent by a large margin across all corpora a search on the lemma *VICTIM* was carried out. WordSketches on *VICTIM* showed that *SURVIVOR* was the strongest “and/or” with LogDice above 10.0 for the Twitter corpora (TSH03\_2019\_06\_Jun LogDice 10.6, TMT03\_2019\_06\_Jun LogDice 10.9). The NSH03\_2019\_06\_Jun corpora for *VICTIM*, had “and/or” matching with *SURVIVOR* at LogDice 9.6, ranking 3<sup>rd</sup> behind *witness* and *perpetrator*. Table 11 illustrates the much higher frequencies of both lemmas in these corpora compared with the general web text enTenTen15<sup>41</sup>.

<sup>39</sup> 22 of the 24 months of data gathering were within the first Trump presidency.

<sup>39</sup> Where a comparison was needed with general internet usage, early in the project the 52 billion word enTenTen15 corpus was used (SketchEngine, 2022), and later the enTenTen21.

<sup>40</sup> LogDice is discussed in section 3.3.8.2.

<sup>41</sup> Where a comparison was needed with general internet usage, early in the project the 52 billion word enTenTen15 corpus was used (SketchEngine, 2022), and later the enTenTen21.

	Frequencies pmw			LogDice
	VICTIM	SURVIVOR	SURVIVORS/VICTIMS	
NSH03_2019_06_Jun	1044.71	79.14	8%	9.6
TSH03_2019_06_Jun	981.43	167.21	17%	10.6
TMT03_2019_06_Jun	831.33	535.33	64%	10.9
enTenTen15	71.5	22.62	32%	8.5

Table 11, Frequencies of VICTIM and SURVIVOR pmw based on 03\_2019\_06Jun data.

This strong relationship across all corpora, and distinction from general web texts, was the basis for choosing VICTIM and SURVIVOR as the focus of the sexual violence aspect of this study.

### 3.3.8 Corpus interrogation (SketchEngine and #LancsBox)

SketchEngine had been used for the early stages of this study because I had been introduced to it earlier in my studies, and hence it presented no further learning curve.

Becoming aware of LancsBox I evaluated it for advantages over SketchEngine particularly intending to include a section on whelks<sup>42</sup> using the 24 monthly corpora to support a diachronic analysis. However, because of the volume of other findings the study's scope had to be reduced.

The corpus interrogation was conducted in SketchEngine as described in section 3.3.8.1.

#### 3.3.8.1 SketchEngine WordSketch and SketchDifference

Within the first few months of data gathering truth had begun to appear as a point of contention for VICTIM but not SURVIVOR. SketchEngine's WordSketch<sup>43</sup> and SketchDifference<sup>44</sup> tools proved to be the most effective way to identify which features of the corpora should be investigated qualitatively.

A WordSketch run in the TwitterSH\_1Yr corpus on the word VICTIM returned a number of modifiers suggesting truth (Table 12). This shows that the words with the four highest LogDice<sup>45</sup> scores and the high frequencies of *real* (9.92), *actual* (9.45), and *true* (8.8) imply the existence of unmarked victims who are in some way bogus. Contestation of truth is discussed further in section 3.4.2.2 and Chapter 6.

<sup>42</sup> Kilgarriff used *whelk* to illustrate the uneven distribution of specialist words within a corpus, whereby a word such as *whelk* might have a high frequency within a single source such as a book on shellfish, but otherwise be rare (1997, p. 138).

<sup>43</sup> WordSketch provides a single page view of a search term's collocates arranged by grammatical function, sorted by frequency or LogDice.

<sup>44</sup> SketchDifference merges the WordSketches for two search terms highlighting words collocating with only one search term, and those common to both (Table 13).

<sup>45</sup> LogDice is discussed in section 3.3.8.2.

Modifiers of VICTIM	Frequency	LogDice Score
<i>real</i> <i>real victims of sexual harassment</i>	31	9.92
<i>rape</i> <i>rape victims</i>	29	9.57
<i>actual</i> <i>actual victims of sexual</i>	18	9.45
<i>true</i> <i>true victims of sexual</i>	10	8.8

Table 12, WordSketch for VICTIM within the TwitterSH\_1Yr corpus

SketchDifference merges the results of two WordSketches providing an easy way to compare differences and similarities in the use of the two lemmas, as noted by Taylor and Marchi (2018, p. 42). For example, Table 13 shows the highest LogDice rated modifiers applied exclusively to VICTIM (in green) or SURVIVOR (in red) and those that are applied to both (in white). The modifiers *real*, *actual*, *alleged*, *true* and *female* only appear with VICTIM as subject, while *CSA*<sup>46</sup>, *student*, *adult*, *care* and *#metoo* only appear with SURVIVOR as subject.

<sup>46</sup> CSA, child sexual assault

Grammar relation	VICTIM frequency (Total 1,769)	SURVIVOR frequency (Total 260)	VICTIM LogDice Score	SURVIVOR LogDice Score
<i>real</i>	29	0	9.6	0
<i>actual</i>	19	0	9.3	0
<i>alleged</i>	15	0	8.9	0
<i>true</i>	10	0	8.5	0
<i>female</i>	11	0	7.9	0
<i>rape</i>	40	6	9.8	7.9
<i>assault</i>	31	7	6.4	8.2
<i>male</i>	9	2	7.9	7.0
<i>child</i>	5	2	7.3	7.4
<i>violence</i>	4	2	7.1	7.7
<i>abuse</i>	5	6	7.2	8.9
<i>#metoo</i>	0	2	0	7.9
<i>care</i>	0	2	0	8.9
<i>adult</i>	0	2	0	9.0
<i>student</i>	0	4	0	9.2
<i>CSA</i>	0	2	0	9.2

Table 13, SketchDifference for VICTIM and SURVIVOR in the TwitterSH\_1Yr corpus

The top ranking modifiers exclusive to VICTIM (green), SURVIVOR (red), and the strongest modifiers shared by both (white).

The difference in frequencies of VICTIM (1,769) and SURVIVOR (260) (Table 13) within the corpus indicates that the numbers must be read with caution, but could suggest different patterns of language use around two words of interest.

### 3.3.8.2 Statistics

SketchEngine provides a number of statistical measures and these were assessed to find the most useful option for this study. Using the word VICTIM within the TwitterMT\_1Yr corpus the top ten collocates for each measure were generated (Table 14). Log Likelihood and T-score both highlight prepositions, punctuation, and definite and indefinite articles, which resulted in less useful sets of concordances for this study. The MI concordances tended to foreground celebrities and politicians rather than representations of the victimised. LogDice which “identifies combinations of words which appear both uniquely in each other’s company but also frequently in the discourse” (Brezina, 2018, p. 274) resulted in collocates more salient to identifying representations of individuals in this study. Hence, LogDice was chosen as the statistical typicality measure.

Collocate of VICTIM (TOTAL 3,439)	LogDice	Collocate of VICTIM	Log likelihood	Collocate of VICTIM	MI	Collocate of VICTIM	T-score
<i>real</i>	10.27179	<i>of</i>	2,129.93	<i>Oldcorn</i>	9.68555	<i>of</i>	20.81
<i>sexual</i>	9.92378	<i>the</i>	1,564.40	<i>crumbling</i>	9.68555	<i>the</i>	19.73
<i>rape</i>	9.74984	.	1,085.88	<i>coverage</i>	9.36362	.	18.09
<i>assault</i>	9.7055	<i>a</i>	1,047.98	<i>Quinn's</i>	9.32298	<i>a</i>	16.10
<i>abuse</i>	9.67079	<i>sexual</i>	804.30	<i>Zoe</i>	8.90794	<i>to</i>	14.50
<i>of</i>	9.57226	<i>real</i>	768.45	<i>AGAIN</i>	8.81882	<i>and</i>	13.85
<i>are</i>	9.34012	<i>to</i>	734.00	<i>Christopher</i>	8.81108	,	13.07
<i>blaming</i>	9.1784	<i>and</i>	717.79	<i>blaming</i>	8.42921	#	12.75
<i>AGAIN</i>	9.11305	<i>are</i>	703.63	<i>husband's</i>	8.42251	<i>are</i>	12.17
<i>believe</i>	9.07	<i>rape</i>	521.90	<i>Professional</i>	8.36362	<i>sexual</i>	11.71

Table 14, Comparison of top 10 tokens for MI, LogDice, T-score and Log-likelihood in the TwitterMT\_1Yr corpus (SketchEngine, VICTIM concordances, showing collocations).

The SketchEngine formulae for calculating LogDice is reproduced here as Figure 14.

The LogDice score indicates the strength of a collocate + lemma pairing. This score is comparable between corpora.

Figure 14, LogDice calculations used in SketchEngine WordSketches

**Dice**

$$\text{Dice}(f_A, f_B) = \frac{2 \frac{f_A}{N} \frac{f_B}{N}}{\frac{f_A}{N} + \frac{f_B}{N}} \simeq \frac{2 \frac{f_{AB}}{N}}{\frac{f_A}{N} + \frac{f_B}{N}} = \frac{2f_{AB}}{f_A + f_B}$$

**logDice**

$$14 + \log_2 \text{Dice} \left( \frac{\|w_1, R, w_2\|}{\|w_1, R, *\|}, \frac{\|w_1, R, w_2\|}{\|*, *, w_2\|} \right) = 14 + \log_2 \frac{2 \cdot \|w_1, R, w_2\|}{\|w_1, R, *\| + \|*, *, w_2\|}$$

<https://www.sketchengine.eu/wp-content/uploads/ske-statistics.pdf>

*N* = corpus size, *f<sub>A</sub>* = number of occurrences of the keyword in the whole corpus (size of concordance)  
*f<sub>B</sub>* = number of occurrences of the collocate in the concordance,  
*f<sub>AB</sub>* = number of occurrences of the collocate in the concordance (co-occurrences),  
*R* = gramrel,  
*w<sub>1</sub>* = word<sub>1</sub>

### 3.4 Concordance categorisation

Categorisation was complicated by the number of concordances using implicit language, requiring an understanding of the wider text or the social environment. Section 3.4.1 describes how this led to the selection of van Leeuwen's (2008) social actor representations and legitimisation frameworks. Particularly, van Leeuwen's Excluded social actors proved useful in identifying individuals who are not directly referenced. These included the Backgrounded individuals who are mentioned elsewhere in the text, and the Suppressed who are entirely absent.

In addition to the van Leeuwen social actor codes and legitimisation types two three-way flags were created: one to provide high level indicators of negative and positive prosody (from

here forward called +/- prosody); and another for veracity. These four variables relate concordances to the research questions (Table 15).

Variables	Possible values	Research questions
Van Leeuwen's (2008) Social Actor Type codes	Listed in section 3.4.1	RQ2 and RQ3
Van Leeuwen's forms of legitimation	Listed in section 3.4.1	RQ2, RQ3, and RQ4
+/- Prosody (three-way flag)	Positive, negative and neutral	RQ1 and RQ3
Veracity (three-way flag)	true, untrue and neutral	RQ2 and RQ3

*Table 15, Variables, categories, and research questions*

Concordances may have more than one Van Leeuwen code and all codes that applied were coded to them. Only single values were assigned for the +/- prosody and veracity 3-way flags.

When categorising multi-phrase tweets the entire tweet was assessed rather than just the KWIC phrase and for Nexis articles the surrounding text was checked to make sure the search term wasn't being used in a way that would change the meaning, for example sarcastically.

Data was prepared for coding by carrying out a Concordance search using the KWIC VICTIM, SURVIVOR or PAY GAP followed by a WordSketch Shuffle<sup>47</sup> on the full 24 month corpora<sup>48</sup> and then downloading the corpus to be opened in Excel. Concordances were coded in batches of 100 until data saturation was reached following the iterative process outlined by Sinclair (1991, p. 84). Data saturation was defined as being reached when no new van Leeuwen social actor or legitimation type had been discovered in the last batch of 100 for any of the corpora. This yielded 500 coded concordances for each corpus/KWIC combination that from here forward are referred to as the vL500 datasets<sup>49</sup> because they were analysed manually from that point.

Tweets that were difficult to assign to a category were revisited over a few days and the meaning then usually became clearer. Where occasional tweets still did not make any sense perhaps because they were a few disjointed words from a single turn in a conversation those concordances were marked as N/A and removed from the sample and replaced with the next concordance in the shuffled sequence. While every effort was made to identify at least one van Leeuwen code for each concordance if none applied then the concordance remained in the sample with only the +/- prosody and veracity flags.

The coding process is now described in more detail, first for van Leeuwen social actor types in section 3.4.1 and then for veracity and +/- prosody in section 3.4.2.

<sup>47</sup> WordSketch Shuffle is a randomisation feature.

<sup>48</sup> TwitterSH, TwitterMT, TwitterPG, NexisSH, NexisMT, and NexisPG.

<sup>49</sup> The ten datasets are listed in Appendix B.

### 3.4.1 Van Leeuwen social actor and authority legitimization types

Van Leeuwen, a socio-linguist, states that "...all texts, all representations of the world and what is going on in it, however abstract, should be interpreted as representations of social practices." (2008, p. 5). This being the case, the language within the corpora can be taken as a reflection of what individuals are doing, saying and believing in their everyday lives.

Representations from van Leeuwen's Social Actor framework (2008, pp. 23-54) provide a practical tool for categorising and discussing a wide range of actors within the text. Van Leeuwen (2008, pp. 23-54) identifies 50 social actor types, of which 29 were identified at least once in the corpora: Abstraction, Activation, Aggregation, Appraisal, Association, Backgrounding, Beneficialisation, Circumstantialisation, Classification, Collectivisation, Connotation, Determination, Differentiation, Dissociation, Formalisation, Functionalisation, Genericisation, Honorific, Indetermination, Indifferentiation, Informalisation, Inversion, Objectivication, Participation, Possessivation, Relational, Semiformalisation, Subjection, and Suppression. This framework has the scope to capture individuals who are either missing from the text (Suppressed) or referred to elsewhere (Backgrounded). The absence of the Suppressed actors may indicate a stylistic choice but alternatively could be an innocent omission (Partington, 2014, p. 135). Partington (2014, p. 136) found that even when unable to articulate their reasoning, people were generally able to interpret media prosody fairly reliably.

The 500 concordances of each dataset were analysed and the lowest level social actor and legitimization type that could be identified in the network hierarchies were assigned. The social actor types reaching the threshold for inclusion are analysed in Chapters 4 and 5 and these are described here using examples from van Leeuwen's analysis of the "Our Race Odyssey" article on immigration from the Sydney Morning Herald, 12 May 1990:

Abstraction	Evaluation is achieved by representing social actors, e.g. the "poor, black, unskilled, Muslim, or illegal" as the conceptual "unwanted problems" with which the group is then identified. P 46
Activation	The actor is an active and dynamic agent of a verb, e.g. "young white thugs attacked African street vendors". P33
Appraisal	The actor is evaluated as "good or bad, loved or hated, admired or pitied", e.g. "young white thugs attacked African street vendors". P 45
Backgrounding	The actor is excluded, but their actions or acts upon them remain as a trace, e.g. "the killing of demonstrators". P 29
Beneficialisation	The actor is a 3 <sup>rd</sup> party and passivated by receiving something, either positive or negative, e.g. "Australia was bringing in about 70,000 migrants a year." P33
Indetermination	The actor remains unspecified and anonymous often using indefinite pronouns such as "some" or "somebody" and can be aggregated by the use of plurals, for example "many believe" P 39, 40.
Indifferentiation	The actor is explicitly positioned as similar to another to show alignment and support, e.g. using "our" and "us". P 40

Subjection            The actor is passivated by reduction to an object, e.g. by nominalising in “An intake of some 54,000 skilled immigrants is expected this year.”. P34

Suppression        The actor is excluded with no trace in the text, e.g. “concerns are being expressed”. P 29

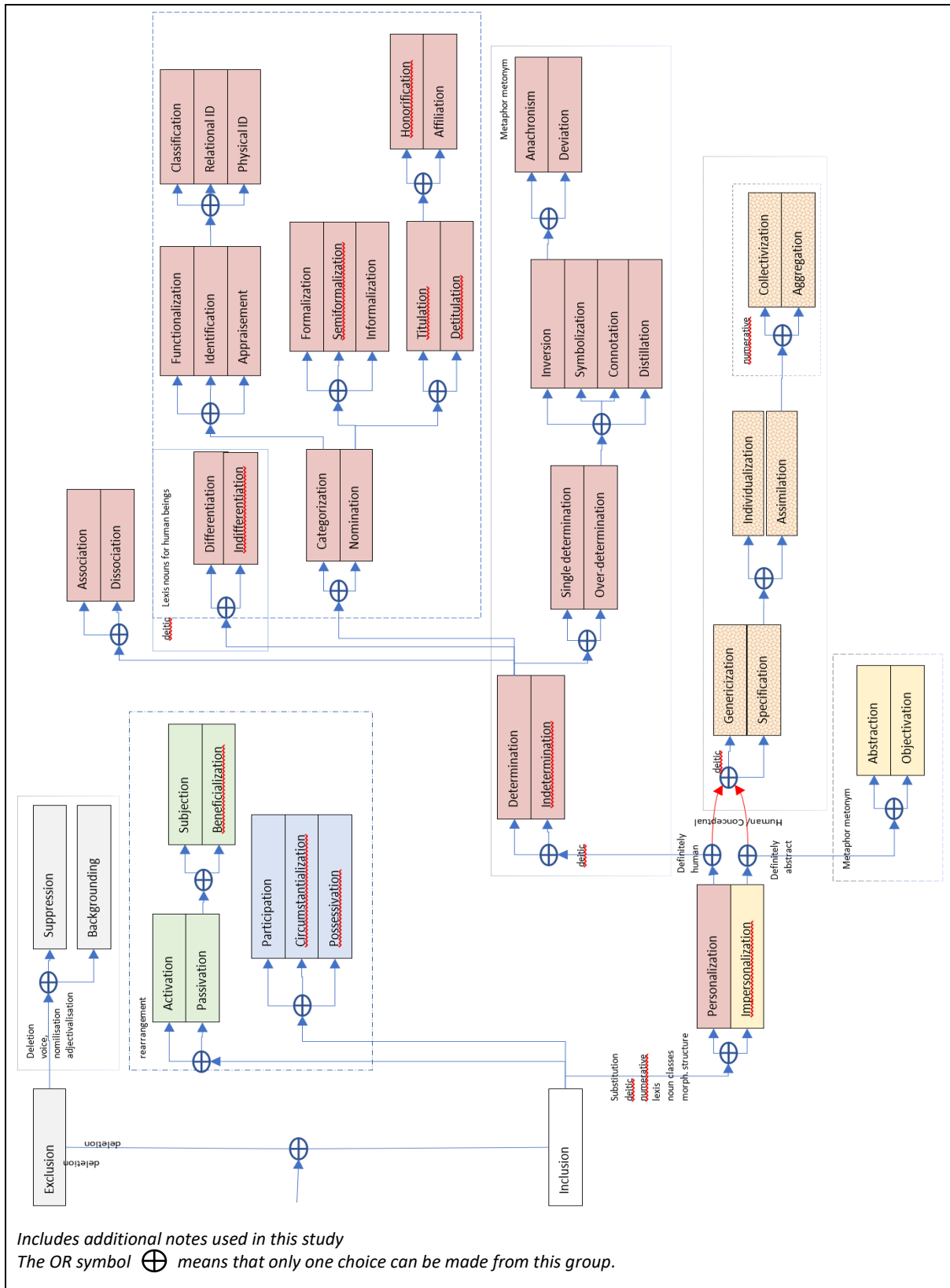
At the top level van Leeuwen’s (2008) social actors are divided into Exclusion and Inclusion categories and beneath these are branching sub-divisions of social actor types. Van Leeuwen’s (2008, p. 52) Social Actor Network diagram is redrawn<sup>50</sup> as Figure 15 . Some social actor representations have an exclusive relationship such as Suppression OR Backgrounding, while others can co-occur, for example Activation AND Personalisation. With multiple social actor representations applying to some concordances the total number sometimes exceeds the number of concordances in a sample. The hierarchical structure means that lower level representations inherit the properties of those above them. The lowest level social actor type that can be identified is used in these datasets.

The first 100 concordances coded from each corpus were used for a pilot analysis, described in Chapter 4.

---

<sup>50</sup> The diagram was redrawn to provide a graphic of higher resolution when expanded to a more readable size than the original. Additionally, the researcher’s annotations were incorporated as an aide memoir.

Figure 15, Re-drawing of van Leeuwen's (2008) Social Actor Network



Particularly useful for this study is van Leeuwen's model of authority legitimation that can be applied to discourse to reveal the authority of social actors (2008, pp. 105-123). *Legitimation* can be used to strengthen authority, but also undermined in *delegitimation*. In *Discourse and Practice*, van Leeuwen (2008) largely discusses legitimation, and a search of the electronic version returns only five instances of the word *delegitimation*, always alongside either *or* or *also*, suggesting that legitimation can be used in a positive way (to legitimise) or negative way (to delegitimise) (2008, p. 106). Legitimation and delegitimation can be implicit and invisible in a tight reading of the text, suggesting a meaning so deeply ideological that it is considered too obvious to need to be articulated (Van Leeuwen, 2008, p. 20).

Van Leeuwen (2008, pp. 105-123) provides four high level categories: Authorisation which would answer the question of "why we should do this", or "why this way" (van Leeuwen, 2008, p. 106); Moral evaluation which is based on moral values that may be subjective or merely implied such as "good or bad", "normal" or "natural" (van Leeuwen, 2008, pp. 109, 110); Rationalisation which refers to "goals, uses, and effects", or make an explicit call to a "natural order of things" (van Leeuwen, 2008, p. 113) ; and Mythopoesis which operates through storytelling to convey a message (van Leeuwen, 2008, pp. 117, 118). This study approached delegitimation via SketchEngine's evaluative adjective predicates and "pay gap" is a ... for the lemma PAY GAP (section 6.2). Concordances with evaluative terms as a KWIC were analysed and coded with van Leeuwen's legitimation types which will be discussed in Chapter 6, and are described here:

#### **Authorisation**

Personal	Authority an individual has through their role and relative status such as a policeman, or a mother to a child, e.g. "because I say so". P 106
Expert	Authority gained through expertise, e.g. "some experts say...". P 107
Role Model	Authority of an admired peer or celebrity to influence others, e.g. "the wise teacher finds out...". P 107
Tradition	Authority gained from longstanding practice, e.g. "this is what we have always done". P 108
Conformity	Authority gained from common practice, e.g. "this is what everyone does". P 109

#### **Moral evaluation**

Evaluation	Evaluation represents a person or action in a way that is "covert and seeks to shield it from debate or argument", e.g. by positioning something as natural or normal. PP 110, 111
Abstraction	Evaluation is achieved by a linkage to a discourse of positive moral values, e.g. "the child goes to school" becomes "the child takes up independence". P 111
Analogy	Evaluation is by making a comparison with a familiar example associated with the required values, e.g. "... as automatic as cleaning your teeth". P 112

## Rationalisation

Theoretical            Legitimation that excludes moral judgement and sense of purpose, but justifying with a definition, explanation, or prediction of “the way things are” akin to naturalization. This study also identifies examples of the Rational Theoretical subtype Experiential which can appear as commonsense knowledge, or proverbs. PP 115, 116

Van Leeuwen’s focus was on legitimation, however there is some research on delegitimation. A Lancaster library search of the two terms *delegitimation* and *van Leeuwen* in any field, in the English language, returned 10 articles. All were published in the last five years and most based on traditional news media or speeches, however two refer to on-line data sources. Both of these find that hegemonic ideologies were able to defend themselves effectively from challenges made by advocates of minority interests suggestive of the reaction against #MeToo (section 4.5).

In the first, Wong (2025) identifies the characteristic styles of delegitimation used in a debate between Singapore’s Ministry of Home Affairs (MHA) press releases and Richard Branson’s blog postings, regarding capital punishment for drug trafficking. Wong found the MHA’s strong use of van Leeuwen’s legitimation strategies presented a “serious challenge” to human rights advocates (Wong, 2025, p. 15).

The second study based on Sino Weibo, a Chinese social media platform, used van Leeuwen’s legitimation types to analyse portrayals of homosexuality (Zhang & Zhuang, 2023, p. 1). They found that the public writing on Weibo used established sociocultural themes such as “nationalism, collectivism, and patriarchy” that promoted delegitimising anti-homosexual representations that were in line with the prevailing ideology (Zhang & Zhuang, 2023, pp. 7, 8).

While van Leeuwen’s analyses are clear and helpful, his examples are from formal written text, which contrasts with the sometimes incomplete, grammatically incorrect, and messy real-world concordances of the present study. Twitter concordances especially, were much more ambiguous and difficult to categorise than his descriptions anticipate. Bouvier (2022, p. 180) described this messiness as “chaos chambers” encompassing a variety of discourses, that existed alongside the more often described “echo chambers” of similar discourses. She related these to a shift in the civil/uncivil boundary in social and political discussions, that contrasts with van Leeuwen’s “clear ‘discursive scripts’”.

Halliday (2014) describes social actors in relation to their grammatical context and his Functional Grammar was initially considered as a tool for this study. However, it appeared less suited to identifying the missing when compared to van Leeuwen’s Social Actor Network and also lacked tools for addressing legitimation, and accommodation of a wider reading of the social environment.

### 3.4.2 Veracity and +/- prosody flags

The vL500 datasets were also coded with +/- prosody and veracity flags. Writers may use semantic prosody to convey meaning that lies along a continuum from the bitterly negative to the glowingly positive. However, for purposes of practicality the coding for this study flattened meaning to just three distinct states. The test for applying the flags was to imagine a reader who has personal experience of the topic being discussed and speculate how they might react. If reading the concordance was likely to reassure and encourage them to share their own story a positive flag was assigned, or a negative flag if they were likely to be discouraged from speaking out. Concordances that felt equally balanced or were neutral in tone were coded as neutral.

The following examples illustrate +/- prosody coding:

**Negative,** @KamalaHarris Her pay gap is a LIE . She’s an embarrassment to women, as are you. TwitterPG\_1Yr

**Positive,** ... hearings on workplace sexual harassment and used the testimony we heard from brave victims, advocates & experts to craft and pass critical protections... TwitterSH\_1Yr

**Neutral,** “Everyone is legally entitled to equal pay for equal work, but there may be a range of factors behind pay gaps and it doesn’t necessarily mean that unlawful discrimination is happening,” says a spokesperson for the EHRC. NexisPG\_1Yr

Similarly the coding for veracity is coded to the three states: true, untrue, and neutral. Veracity was deduced from both the explicit expressions of truth such as *myth* and *playing the victim* and also from implicit references to the wider social context. Here the neutral flag includes all concordances that indicate neither truth nor untruth and also those balancing arguments for both possibilities. The three veracity states are illustrated in the following examples:

**True,** I believe that the woman in Gillibrand’s office was subject to sexual harassment & retaliation. I believe that Al Franken sexually harassed & treated women like things... TwitterSH

**Untrue,** The pay gap is a myth. TwitterPG

**Neutral,** Debating the ‘gender pay gap’. Full video here: <https://t.co/Dcl7pg2MqK>  
<https://t.co/sOITzSJPrZ> TwitterPG

Truth was represented differently within the sexual violence and pay gap corpora and this is explored further in sections 3.4.2.1 and 3.4.2.2 respectively.

#### 3.4.2.1 Veracity and sexual violence

Once the \_1Yr corpora were available the frequencies of collocates used in disputing the truth (*false, fake, debunk, believe, myth, and lie*) were obtained from WordSketch Concordance Searches (Table 16). Higher frequencies of words disputing truth were found in the Twitter\_1Yr sexual violence corpora than either the enTenTen15 or NexisSH\_1Yr. The highest frequencies above enTenTen15 levels are for *fake* (16 times higher) and *false* (also 16 times higher) both within the TwitterMT\_1Yr corpus.

	NexisSH_1Yr pmw	NexisMT_1Yr pmw	TwitterSH_1Yr pmw	TwitterMT_1Yr pmw	enTenTen15 pmw
<i>believe</i>	346.41	477.17	480.93	899.14	240.07
<i>debunk</i>	2.49	1.97	8.94	8.37	1.82
<i>fake</i>	16.17	48.05	173.72	252.23	15.8
<i>FALSE</i>	146.15	135.58	206.29	311.4	31.0
<i>lie</i>	16.17 *	147.43	480.93	456.05	84.38*
<i>myth</i>	9.95	36.86	13.41	26.3	18.61
Total Words	1,360,468	1,262,017	1,242,624	1,233,750	13,190 M

Table 16, Frequencies of words used to dispute veracity in sexual harassment

\* Lie within the enTenTen15 is used in the sense of untruth <3%, while in the NexisSH\_1Yr untruths makes up 55%. All instances within TwitterSH\_1Yr describe untruth.

Because the pay gap is discussed as a phenomena rather than as impacting individuals it made sense to approach it slightly differently and that is discussed in the next section.

### 3.4.2.2 Veracity and the pay gap

Representations within the pay gap corpora were approached via the “... is a” construct using a WordSketch of the lemma PAY GAP (Table 17). For TwitterPG\_1Yr more varied words were used in the “... is a” construct but only the truthfulness ones are listed in the table. In contrast, for NexisPG\_1Yr all are displayed none of which reference the truth.

TwitterPG_1Yr			NexisPG_1Yr		
Collocate	Freq	LogDice	Collocate	Freq	LogDice
<i>myth</i>	392	12.8	<i>difference</i>	23	12.2
<i>lie</i>	31	9.8	<i>issue</i>	26	12.0
<i>bullshit</i>	19	9.1	<i>result</i>	7	10.3
<i>hoax</i>	8	7.9	<i>something</i>	6	10.3
<i>reality</i>	7	7.7	<i>problem</i>	4	9.9

Table 17, “PAY GAP is a...” collocates from a pay gap WordSketch

The contrast between the strong presence of truth related words within TwitterPG\_1Yr against their absence from NexisPG\_1Yr is striking. The theme of disputing truth in both the sexual violence and pay gap corpora are discussed further in analysis Chapters 4 to 6.

## 3.5 Identifying the themes

This section describes the capture of the main themes analysed in Chapters 4 to 6. For the purpose of the present study *theme* is defined as a group of representations sharing a common feature and for VICTIM and SURVIVOR these are of two types:

- Those that map cleanly to the van Leeuwen social actor codes of the same name: Appraisalment, Indifferentiation, and Passivation;
- Those identified and named within this study: Abuse; Blaming/shaming; Causes; Consequences; Definitions; Denial; Playing (the victim card); Speaking out; and Victim vs survivor.

For the pay gap a trio of high level themes: Veracity, Acceptability, and Contempt emerged with subthemes similar to those within the sexual violence datasets organised below them. These sub-themes either map to van Leeuwen social actors or were identified within the data:

- Appraisalment, Indifferentiation, Passivation, and Suppression;
- Abuse, Causes, Consequences, Definitions, Denial, Speaking out.

The pay gap texts frequently focused on discussions about the statistical phenomenon of the pay gap; while texts involving sexual violence generally focus on individuals or events. Consequently, while the representations of the targets of SEXUAL HARASSMENT could be accessed directly as the objects and subjects of verbs, those disadvantaged by the pay gap had to be approached indirectly through representations of the pay gap itself. In this way accessing the most prominent patterns within the corpora allowed the similarities and contrasts to be identified for a deeper level of qualitative analysis.

Section 3.5.1 describes the route taken through the sexual violence corpora, starting with the verb collocates of VICTIM and SURVIVOR as object and subject. The path taken through the pay gap corpora using the “pay gap is a...” construct is outlined in section 3.4.2.2 before being contrast with the findings from the sexual violence corpora. The missing and muted are addressed in section 3.5.2. Section 3.5.3 completes the section with a brief discussion of how agency manifests in these corpora.

**3.5.1 Victims and survivors**

While VICTIM is more common in all corpora SURVIVOR is preferred by some writers especially within #MeToo tweets, some of which explicitly discuss this choice of word to which they attribute more agency than VICTIM. SURVIVOR also implies an event in the past with a sense of moving-on or being “over it” while VICTIM suggests ongoing issues suggesting an earlier stage of healing. However the agency that SURVIVOR carries is complicated because of the connotation of personal responsibility, and that will be discussed in Chapter 7.

The following analysis was carried out at an early stage so is based on the \_03\_2019\_06\_Jun corpora. Focusing first on the sexual harassment and #MeToo corpora, SURVIVOR appeared much less frequently than VICTIM in all of the \_03\_2019\_06\_Jun corpora (Table 18).

pmw	VICTIM pmw words (VICTIM/VICTIM + SURVIVOR)	SURVIVOR pmw words SURVIVOR/VICTIM + SURVIVOR)	VICTIM + SURVIVOR
TMT_03_2019_06_Jun	1,214.39 (65%)	657.02 (35%)	1871.41
TSH_03_2019_06_Jun	1,025.93 (84%)	201.49 (16%)	1227.42
NMT_03_2019_06_Jun	631.9 (75%)	210.63 (25%)	842.53
NSH_03_2019_06_Jun	802.66 (86%)	130.51 (14%)	933.17
BNC	59.2 (86%)	9.53 (14%)	68.73
EnTenTen15	68.97 (76%)	21.63 (24%)	90.6

*Table 18, Frequency of the lemmas VICTIM and SURVIVOR pmw within the \_03\_2019\_06\_Jun corpora Percentages show the lemma’s part of the combined VICTIM plus SURVIVOR total.*

Looking at the sum of VICTIM plus SURVIVOR there are dramatic differences with TMT\_03\_2019\_06\_Jun having the highest frequencies and with half that number

NSH\_03\_2019\_06\_Jun has the lowest. Within each corpora the proportion of frequencies for SURVIVOR and VICTIM also varied between TMT\_03\_2019\_06\_Jun with the highest percentage of SURVIVOR (35%) and NSH\_03\_2019\_06\_Jun the lowest (14%). The frequencies were so varied that the EnTenTen15 and BNC corpora were added to provide a sanity check. The BNC returned the same percentage split between the two terms VICTIM and SURVIVOR as NSH\_03\_2019\_06\_Jun contrasting with the EnTenTen15 corpus which was midway between TMT\_03\_2019\_06\_Jun and TSH\_03\_2019\_06\_Jun. However these corpora have frequencies that were orders of magnitude higher than either the EnTenTen15 or the BNC suggesting different patterns of use. The following sections explore the main themes within analyses of the larger \_1Yr corpora.

### **3.5.1.1 VICTIM and SURVIVOR as subject and object of verbs**

Once the first 12 monthly data files had been prepared they were aggregated into \_1Yr corpora to provide a broader overall picture. Using the TwitterSH\_1Yr corpora a SketchEngine SketchDifference was run using the lemmas VICTIM and SURVIVOR and this showed interesting contrasts and similarities in the uses of their collocates (Table 19). For instance *blame* collocates only with VICTIM in all corpora, and *play* (as in *playing the victim*) occurs only in the Twitter corpora. There is no *playing the survivor*. Consequently VICTIM and SURVIVOR as the object and subject of verbs proved to be a fruitful area for this study, discussed further in Chapter 4.

A surprising discovery was *silence* as the object of a verb in the TwitterSH\_1Yr corpus for the VICTIM lemma. Having seen the theme of truth and legitimacy emerge within the pay gap corpora this raised the question of whether the silencing of certain voices articulating sexual violence might be part of the same theme. *Silence* itself did not occur within WordSketches of either of the pay gap corpora but may manifest as an activity for instance through the delegitimising of individuals. The lemma SILENCE is discussed further in section 3.5.2.

Sexual violence and the gender pay gap

3 Methodology

		VICTIM	Frequency/LogDice	SURVIVOR	Frequency/LogDice
verbs with VICTIM/SURVIVOR as object	NexisSH_1Yr	<i>compensate</i>	17, 10.2	<i>stand</i>	1, 9.7
		<i>help</i>	12, 9.3	<i>let</i>	4, 9.5
		<i>blame</i>	10, 8.9	<i>injure</i>	1, 9.5
		<i>urge</i>	6, 8.8	<i>shame</i>	1, 9.4
				<i>re-traumatize</i>	1, 9.4
	NexisMT_1Yr	<i>silence</i>	15, 10.1	<i>help</i>	10, 10.2
		<i>help</i>	17, 9.9	<i>support</i>	11, 10.1
		<i>blame</i>	9, 9.4	<i>silence</i>	4, 9.7
		<i>encourage</i>	10, 9.3	<i>enable</i>	2, 8.9
		<i>support</i>	8, 8.7	<i>empower</i>	2, 8.7
		<i>believe</i>	6, 8.4		
	TwitterSH_1Yr	<i>blame</i>	47, 10.9	<i>Inspire</i>	9, 11.3
<i>silence</i>		27, 10.2	<i>fail</i>	7, 10.7	
<i>play</i>		12, 8.9	<i>shoot</i>	3, 9.7	
<i>fail</i>		10, 9.38.3	<i>select</i>	2, 9.4	
<i>believe</i>		12, 8.7	<i>empower</i>	2, 9.1	
<i>attack</i>		9, 8.6	<i>ensure</i>	2, 8.8	
TwitterMT_1Yr	<i>blame</i>	25, 9.5	<i>fail</i>	3, 8.6	
	<i>play</i>	25, 8.6	<i>meet</i>	3, 8.4	
	<i>become</i>	12, 8.4	<i>mean</i>	3, 8.0	
	<i>shame</i>	8, 8.0			
	<i>claim</i>	4, 8.0			
verbs with VICTIM/SURVIVOR as subject	NexisSH_1Yr	<i>fear</i>	16, 10.5	<i>navigate</i>	2, 10.6
		<i>corroborate</i>	6, 9.2	<i>differ</i>	1, 9.5
		<i>feel</i>	7, 8.9	<i>like</i>	1, 9
		<i>blamed</i>	3, 8.7	<i>sit</i>	1, 8.3
		<i>inform</i>	1, 8.4	<i>experience</i>	1, 7.4
	NexisMT_1Yr	<i>blame</i>	12, 10.1	<i>testify</i>	5, 9.6
		<i>suffer</i>	8, 9.3	<i>strive</i>	2, 9.1
		<i>speak</i>	9, 8.9	<i>speak</i>	6, 8.9
		<i>shame</i>	5, 8.8	<i>encourage</i>	2, 8.8
		<i>break</i>	4, 8.1	<i>share</i>	3, 8.7
		<i>lose</i>	3, 7.9		
	TwitterSH_1Yr	<i>blame</i>	41, 12.0	<i>share</i>	2, 8.9
		<i>shame</i>	5, 9.5	<i>feel</i>	2, 7.8
		<i>suffer</i>	4, 8.6	<i>speak</i>	2, 7.6
<i>report</i>		6, 8.4	<i>think</i>	2, 6.9	
<i>need</i>		7, 8.3	<i>get</i>	2, 5.7	
TwitterMT_1Yr	<i>blame</i>	28, 10.7	<i>tell</i>	7, 8.8	
	<i>shame</i>	19, 9.7	<i>share</i>	6, 8.7	
	<i>feel</i>	7, 9.4	<i>know</i>	5, 8.5	
	<i>speak</i>	12, 9.1			
	<i>get</i>	20, 8.7			

Table 19, Verbs with VICTIM and SURVIVOR as object and subject in the \_1Yr corpora

Within Table 19 *blame* and *shame* collocate with VICTIM and SURVIVOR however, there is no similar noun for those disadvantaged by the pay gap for *blame* or *shame* to collocate with, and SketchEngine WordSketches found no collocations with *woman* or *employee* within the pay gap corpora. This does not mean that *blaming* or *shaming* does not exist within the pay gap corpora, rather that *blaming* and *shaming* are less often discussed explicitly. Explicit victim-blaming in relation to the sexual violence corpora is discussed again in Chapter 4.

### 3.5.1.2 *Blaming and shaming*

Frequencies of BLAME and SHAME are highest within TwitterSH\_1Yr (185.99 and 171.78 respectively) and TwitterMT\_1Yr (191.64 and 238.85 respectively) (Table 20).

	NexisSH_1Yr freq, pmw	NexisMT_1Yr freq, pmw	NexisPG_1Yr freq, pmw	TwitterSH_1Yr freq, pmw	TwitterMT_1Yr freq, pmw	TwitterPG_1Y freq, pmw
BLAME	108 (65.1)	142 (98.82)	100 (66.9)	288 (185.99)	276 (191.64)	118 (75.00)
SHAME	78 (47.02)	202 (140.58)	68 (45.49)	266 (171.78)	344 (238.85)	142 (90.26)

Table 20, Frequency in pmw of *blame* and *shame* in \_1Yr corpora

To test the strength of the BLAME and SHAME collocate pairings with *victim* and *survivor* four further WordSketches were run using the verb lemmas BLAME and SHAME as the object and subject of *victim/survivor* across the \_1Yr corpora shown in: Table 21 (BLAME + *victim*), Table 22 (SHAME + *victim*), Table 23, (BLAME + *survivor*), and Table 24 (SHAME + *survivor*). Where SHAME is found at all, it always scores lower than BLAME. This is significant because BLAME is a causal statement (Oxford English Dictionary [electronic resource], n.d.-e) casting the individual as culpable and less worthy of sympathy, whereas *shame* is an uncomfortable emotional response associated with loss of respect (Oxford English Dictionary [electronic resource], n.d.-E).

BLAME + <i>victim</i>	NexisSH_1Yr	NexisMT_1Yr	TwitterSH_1Yr	TwitterMT_1Yr	BNC	EnTenTen15
Ranking <sup>51</sup>	10	7	1	1	3	1
Frequency	9	9	47	18	21	3,463
LogDice	9.2	9.4	10.8	9.5	7.45	7.5

Table 21, BLAME with *victim* as object in \_1Yr corpora

BLAME often occurs as the compound verb *victim-blaming*, for example “... the same people that supported Dr. Ford and now *victim-blaming* Flores ...” (TwitterSH, *victim*, 2019-04, 63). BLAME is seen at high rates in all of the sexual violence corpora but ranks highest in TwitterSH\_1Yr (10.8 LogDice) where *victim* is the top ranking object of BLAME (Table 21).

<sup>51</sup> Ranking relates to the word’s position in the SketchEngine WordSketch results table.

## Sexual violence and the gender pay gap

### 3 Methodology

SHAME + <i>victim</i>	NexisSH_1Yr	NexisMT_1Yr	TwitterSH_1Yr	TwitterMT_1Yr	BNC	EnTenTen15
Ranking <sup>51</sup>	0	32	0	0	0	0
Frequency	0	1	0	0	0	0
LogDice	0	6.3	0	0	0	0

Table 22, SHAME with victim as object in \_1Yr corpora

In Table 22 *victim* only appears as the object of SHAME within the NexisMT\_1Yr corpus (LogDice 6.3) where 31 other words have higher logDice.

BLAME + <i>survivor</i>	NexisSH_1Yr	NexisMT_1Yr	TwitterSH_1Yr	TwitterMT_1Yr	BNC	EnTenTen15
Ranking <sup>51</sup>	0	0	0	0	0	0
Frequency	0	0	0	0	0	0
LogDice	0	0	0	0	0	0

Table 23, BLAME with survivor as subject in \_1Yr corpora

Table 23 shows *survivor* does not collate as the subject of BLAME in the study's corpora or the BNC or EnTenTen15.

SHAME + <i>survivor</i>	NexisSH_1Yr	NexisMT_1Yr	TwitterSH_1Yr	TwitterMT_1Yr	BNC	EnTenTen15
Ranking <sup>51</sup>	0	0	0	0	0	0
Frequency	0	0	0	0	0	0
LogDice	0	0	0	0	0	0

Table 24, SHAME with survivor as subject in \_1Yr corpora

*Survivor* does not collate as the subject of SHAME (Table 24).

Within these corpora *victim* collates with BLAME and SHAME whereas *survivor* does not. The theme of blaming and shaming continues in Chapters 4 to 6.

#### 3.5.2 Silence

Another interesting theme to emerge from the sexual violence corpora was the prevalence of the verb *silence* with VICTIM and SURVIVOR as object (Table 25 and Table 26). *Silence* has been included as a truth theme as it is used to obscure the truth. This analysis (Table 25) showed that *silence* appeared with high LogDice in NexisMT\_1Yr (10.1) and TwitterSH\_1Yr (10.2) .

NexisSH _1Yr	Freq, Score	NexisMT _1Yr	Freq, Score	TwitterSH _1Yr	Freq, Score	TwitterMT _1Yr	Freq, Score
<i>compensate</i>	17, 10.2	<i>silence</i>	15, 10.1	<i>blame</i>	47, 10.9	<i>blame</i>	25, 9.5
<i>help</i>	12, 9.3	<i>help</i>	17, 9.9	<i>silence</i>	27, 10.2	<i>play</i>	25, 8.6
<i>blame</i>	10, 8.9	<i>blame</i>	9, 9.4	<i>play</i>	12, 8.9	<i>become</i>	12, 8.4
<i>urge</i>	6, 8.8	<i>encourage</i>	10, 9.3	<i>fail</i>	10, 8.8	<i>shame</i>	8, 8.4
		<i>support</i>	8, 8.7	<i>believe</i>	12, 8.7	<i>claim</i>	4, 8.0
		<i>believe</i>	6, 8.4	<i>attack</i>	9, 8.6		

Table 25, Verbs with *victim* as object within sexual violence \_1Yr corpora

It was possible that there were other high scoring objects of the verb *silence* which would make its use in the context of *victim* and *survivor* unexceptional. To rule out this possibility the lemma *SILENCE* was used to drive WordSketches within each of the sexual violence corpora to identify other objects. Table 26 illustrates the association of *SILENCE* with *victim* and *survivor*. While *victim* appears in the top ten ranking verbs of all four sexual violence corpora, *survivor* appears only for NexisMT\_1Yr. *Whistleblower* which appeared in NexisMT\_1Yr, TwitterSH\_1Yr and TwitterMT\_1Yr suggested another potential term for the victimised, however closer inspection of the wider text showed that all instances referred to a specific victim of sexual harassment in a US political scandal, rather than referring to the category of the victimised.

NexisSH_1Yr	Freq, Score	NexisMT_1Yr	Freq, Score	TwitterSH_1Yr	Freq, Score	TwitterMT_1Yr	Freq, Score
<i>restriction</i>	2, 10.6	<i>accuser</i>	7, 10.5	<i>victim</i>	27, 10.2	<i>victim</i>	15, 9.4
<i>precondition</i>	2, 10.1	<i>victim</i>	15, 10.1	<i>Hartley</i>	2, 9.7	<i>#HERstory</i>	1, 9.0
<i>cry</i>	1, 9.8	<i>whistleblower</i>	2, 9.7	<i>voice</i>	4, 9.4	<i>Globe</i>	1, 9.0
<i>score</i>	1, 9.5	<i>survivor</i>	4, 9.7	<i>accuser</i>	2, 9.1	<i>protector</i>	1, 9.0
<i>voice</i>	2, 9.2	<i>assistant</i>	2, 9.5	<i>marginalised</i>	1, 8.7	<i>Me</i>	1, 9.0
<i>accuser</i>	1, 8.9	<i>sister</i>	2, 9.2	<i>whistle-blowing</i>	1, 8.7	<i>whistleblower</i>	1, 8.9
<i>News</i>	1, 8.7	<i>agreement</i>	2, 8.9	<i>picadillo</i>	1, 8.7	<i>settlements</i>	1, 8.9
<i>client</i>	1, 8.7	<i>troll</i>	1, 8.8	<i>critic</i>	1, 8.7	<i>rest</i>	1, 8.8
<i>victim</i>	1, 8.7	<i>complainant</i>	1, 8.6	<i>artist</i>	1, 8.5	<i>#women</i>	1, 8.7
<i>conversation</i>	1, 7.7	<i>organisation</i>	1, 8.4	<i>criticism</i>	1, 8.5	<i>twitter</i>	1, 8.7

Table 26, Top ten most frequent objects of the verb *SILENCE* within the sexual violence \_1Yr corpora.

*Silence* occurs most frequently as a verb in the corpora, however the noun form also contained some points of interest with the top ranking term within NexisMT\_1Yr (13.6), TwitterSH\_1Yr (11.7), and TwitterMT\_1Yr (12.1) being *breaker/break* suggesting that breaking silence may be viewed as important (Table 27). A visual check of instances of *breaker/break* often referred to other situations not relevant to the present study such as *breaking news*, or *broke campaign finance laws*. Viewing the NexisMT\_1Yr *breaker* concordances revealed that half referred to discussion of Time Magazine’s 2017 award of a collective “Person of the Year” award.

3 Methodology

NexisSH _1Yr	Freq, Score	NexisMT _1Yr	Freq, Score	TwitterSH _1Yr	Freq, Score	TwitterMT _1Yr	Freq, Score
<i>survivor</i>	1, 8.5	<i>breaker</i>	23, 13.6	<i>breaker</i>	2, 11.7	<i>break</i>	25, 12.1
<i>culture</i>	1, 6.6	<i>#nomoresilence</i>	1, 10.1	<i>critic</i>	1, 10.3	<i>deafen</i>	2, 10.3
<i>woman</i>	1, 4.8	<i>editorial</i>	1, 10.0	<i>cash</i>	1, 9.9	<i>end</i>	6, 10.0
		<i>habit</i>	1, 9.5	<i>account</i>	1, 8.2	<i>brisent</i>	1, 9.4
		<i>today</i>	1, 7.2	<i>victim</i>	3, 7.2	<i>brisaient</i>	1, 9.4
		<i>victim</i>	1, 6.2	<i>worker</i>	1, 6.4	<i>observe</i>	1, 9.3
		<i>year</i>	1, 4.8	<i>fund</i>	1, 5.2	<i>respond</i>	1, 9.1
				<i>woman</i>	2, 5.1	<i>choose</i>	1, 8.5
						<i>buy</i>	1, 8.3

Table 27, Modifiers of the noun SILENCE within the sexual violence \_1Yr corpora

Silence appears to be a strong theme around VICTIM and SURVIVOR and this will be discussed in more depth in Chapter 7.

**3.5.3 Agency**

Agency is not easily identified by a mechanical search as it is often enacted more subtly and it cross-cuts the other analyses. However while carrying out the pilot study agency also appeared explicitly in some concordances for example:

- Fact: The gender pay gap is real. & you can do something about it by *negotiating* for your salary. (TwitterPG\_1Yr)  
In this example women are advised to be agentic by *negotiating*, suggesting a current lack of it.
- Women talking to women about women’s issues is never going to get us very far. We need those in power ? still mostly men ? to *create* the space and opportunity for us to join them..” (NexisPG\_1Yr)  
Here the call for agency is for *men* to be agentic, passivating women by creating space for them.
- ...settlements reveals a veritable ‘cottage industry’ where lawyers benefit from working out settlements for victims, who are *forced to sign* non-disclosure agreements. (NexisSH\_1Yr)  
The victims are doubly passivated here, firstly in having settlements *worked out for them*, and then in being *forced* to sign NDAs.
- ... these allegations put the blame on women that they use sexual favors to get positions & power. (TwitterMT\_1Yr)  
Women are portrayed as agentic in this example, but it is used to blame them for proactively entering into quid pro quo arrangements. This is a reversal which lends the accuser agency but also damages the legitimacy of their accusation.

How the theme of agency is claimed and denied is complex and it runs as a thread through the whole of the present study so will be referred to in most analyses.

### 3.6 Issues

It was a deliberate decision to try to work with the largest practical corpora as any early filtering could exclude potentially interesting features from the data however size brought its own issues. Initial attempts to aggregate the monthly data files (approximately 1,200,000 words each) caused errors in processing during deduplication and caused WordSmith to crash. A safer solution was to carry out data cleaning and deduplication month by month. After twelve months of data gathering the monthly corpora were deduplicated against each other and this process was discussed in sections 3.3.2 and 3.3.4.

The monthly corpora tended to be just over their target 100K words. Hence, to create the \_1Yr corpora of 1.2 M words a small number of words were removed from each monthly corpus. The largest reduction being for NexisPG\_1Yr which was reduced by 15 percent while TwitterSH\_1Yr and TwitterMT\_1Yr were reduced by only three percent each. Table 28 shows the reduction of the word count following deduplication of the \_1Yr corpora. Consequently a small number of concordances used as examples in pilot analyses may no longer exist in the aggregated \_1Yr corpora. Similarly when aggregating \_1Yr and \_2Yr corpora another small number of words was removed to keep the finalised corpora close to 2.4 M words.

	Sum of words in the monthly corpora	Aggregated 1Yr corpora	% difference in word count
TwitterPG_1Yr	1,294,340	1,242,297	4%
TwitterSH_1Yr	1,270,492	1,229,498	3%
TwitterMT_1Yr	1,211,942	1,176,791	3%
NexisPG_1Yr	1,426,906	1,208,825	15%
NexisSH_1Yr	1,411,525	1,269,081	10%
NexisMT_1Yr	1,262,017	1,205,490	4%

Table 28, Loss of words upon aggregation of corpora<sup>52</sup>

It later transpired that a few unusable<sup>53</sup> tweets had eluded the data cleaning described in section 3.3.5, only appearing during later qualitative analyses. These could not be identified and removed without an impractical amount of manual searching. These contribute to the quantitative analyses and potentially dilute text effects but the effect should be negligible within the 2.4 M word corpora. Where these concordances appeared during qualitative processes they were discounted as not containing language relevant to the present study.

Possible ethical concerns were considered. All tweets and Nexis news articles are already in the public domain, however the authors may not have expected their text to be put to the current use. Additionally, particularly within the tweets, there may be mention of third parties who are unaware of their names being used. Consequently, where a name was not a high profile individual I googled them, and if I failed to find a reference to them within UK mainstream news media websites, their name was redacted.

<sup>52</sup> The monthly corpora figures are unavailable for the 1,262,017 word NexisMT\_1Yr corpus.

<sup>53</sup> Unusable due to being made up entirely of unusual characters not included in the filters.

No research is value free (Gillings et al., 2023, p. 46) and this study has elements that are heavily subjective. While I will endeavour to avoid bias in the decisions and evaluations I make in this study my socialisation as a white, middle-aged, female, and UK-based researcher is bound to have some influence. Subjectivity is discussed in more detail in section 2.3.1, 4.1, 4.2, 4.3.1, 4.3.2.1, 5.4.1 and 8.3.

### **3.7 Conclusion of Methodology**

This chapter has explained how the data was processed and prepared for analysis and also how the pilot analyses were carried out in order to finesse the approach for the main part of the study.

The themes that emerged during the pilot study are analysed in the following chapters:

- Chapter 4 Prosody and sexual violence
- Chapter 5 Prosody and the pay gap
- Chapter 6 Delegitimisation and the pay gap.

Note: concordances in the following chapters are identified using the format  
newsmedia<sup>54</sup> name (corpus, KWIK, yyyy-mm, ID number)

---

<sup>54</sup> Newsmedia name only applies to Nexis examples

ID number is the row number that applies only to the manually coded datasets.

## 4 Prosody and sexual violence

### 4.1 Introduction

This chapter establishes a framework for evaluating the first research question which is repeated below for convenience.

RQ1 Within the context of Twitter and Nexis sexual violence datasets, how are VICTIM and SURVIVOR represented in relation to possible positive<sup>55</sup> and negative affective responses? This question is addressed in three parts:

RQ1A, Are VICTIM representations more frequently negative than SURVIVOR representations? (sections 4.2 & 4.3)

RQ1B, What are the dominant themes associated with VICTIM and SURVIVOR? (sections 4.4 & 4.5)

RQ1C, What is gained by claiming the SURVIVOR label? (section 4.6 & 7)

The analysis in this chapter enables the study to validate or disprove RQ1A's claim that VICTIM representations are more often positioned as negative in the corpora than SURVIVOR representations. The chapter then explores the themes associated with VICTIM and SURVIVOR to answer RQ1B. By the end of the chapter the potential advantages gained by adopting the SURVIVOR label will have been identified to provide an answer to RQ1C.

Analysis in this chapter is based on the vLPilot and vL500 sexual violence datasets<sup>56</sup> that are coded first for negative and positive prosody and then for van Leeuwen social actor types.

The subjectivity that interlocutors, including this researcher, bring to a text are a large part of the back-drop to this chapter affecting any reading of the text. Writers use language to tell a story, create drama, or build an argument. Readers may interpret a multiplicity of different social actors/actions that can be woven dynamically into a single concordance. Baxter (2003, p. 73) advises researchers to seek the multiple, contradictory and shifting voices within a discourse while avoiding fixing on a definitive meaning.

Section 4.2 explains the "why" and "how" of the +/- prosody coding, while section 4.3 does the same for coding of themes emerging around the van Leeuwen social actor types. In section 4.4 examples are used to illustrate the dominant positive representations while negative representations are addressed in section 4.5. Attention turns to more specifically evaluating agency in section 4.6. Section 4.7 addresses RQ1A, RQ1B and RQ1C.

Representations of sexual violence are compared with those within the pay gap corpora in section 5.7.

Note: example tweets are reproduced as they were posted, complete with any spelling or grammatical errors.

---

<sup>55</sup> Positive and negative is introduced and defined in section 3.4.

<sup>56</sup> The sexual violence datasets and corpora comprises: TwitterSH, TwitterMT, NexisSH, and NexisMT.

## 4.2 +/- prosody in representations

In real world communications there is a continuum of meaning between the extremes of negative and positive. Here it has been necessary to flatten prosody to just three distinct states: positive, neutral and negative as explained in section 3.4.2.

Settling on the definitions of positive, neutral and negative that would be most useful in answering the research question was more complex than anticipated and the rationale is discussed in section 4.2.2. The distribution of prosody flags across the vLPilot datasets is described in section 4.2.3, while section 4.2.4 outlines the result of third-party +/- prosody cross-checking. Section 4.2.5 describes the necessary exclusion of a small number of concordances.

During the lifecycle of this study sentiment analysis tools started to become readily available. These initially sounded like a potentially fast, consistent, and less subjective way of identifying +/- prosody, however after a brief investigation they proved unsuitable and this is briefly justified in section 4.2.1.

### 4.2.1 Sentiment analysis tools

Sentiment analysis tools use artificial intelligence (AI) algorithms to analyse the emotional tone of a piece of text as positive or negative. Studies have recently started to use sentiment analysis tools that assign positive and negative sentiment, such as described by Alkhamash (2021). However, on trying out tools, such as Formula Bot (2025) it quickly became clear that the tool was unable to distinguish between a negative statement and a sympathetic writer quoting or challenging a negative statement. For instance both Example 1 and Example 2 in section 4.2.2 were classified by Formula Bot as negative while I would have flagged them as positive because they are supportive of SURVIVORS.

In order to use sentiment analysis a more sophisticated product may be needed. However, a more sophisticated AI product may involve a more lengthy learning curve and set up process putting it beyond current resources. As the present study is more concerned with the reader's likely interpretation, assignment of +/- prosody flags was a manual exercise.

### 4.2.2 Defining positive, negative and neutral prosody

For some concordances assigning flags as positive, neutral or negative was more complicated than anticipated. For example, is the sentence in Example 1 negative or positive?

Example 1     The abuse of power in the alleged sexual harassment case above has left the SURVIVOR extremely traumatised.

Malay Mail Online (vLNSH500, SURVIVOR, 2020-04, 70)

A tight reading positions it as negative because this survivor has been subject to an *abuse of power* and is *extremely traumatised* but a wider reading suggests support in calling out the *abuse of power* and recognition of trauma. Example 1 was coded as positive.

In Example 2 a society-wide problem is highlighted that suggests on-going dire consequences for survivors seeking justice. The geographical context is Bangladesh and this article comes from a reputable source positioning these threats as very real. Again this is a negative situation that may deter a victim from speaking out but conversely the calling out and raising

public awareness could be seen as progress and give confidence. This example was also coded as positive.

Example 2 These problems are compounded because Bangladesh has no witness protection law, meaning that SURVIVORS pursuing legal remedies and those willing to testify on their behalf risk serious threats, harassment, and even death, the statement adds. The Financial Express (Bangladesh) (vLNSH500, SURVIVOR, 2020-10, 103)

Example 3 presents a different challenge because at face value it reports an ideal situation with *no complaints from actual VICTIMS*. At the same time there were 300 *inquiries... about accusations*, which might include questions such as “how many accusations” or “what kind of accusations” had been made, and 300 such questions would indicate a strong expectation of accusations. Extratextual knowledge of a totalitarian regime with a poor human rights record potentially positions it as a suppression of the truth that silences the wronged. Furthermore, denying the existence of sexual harassment may leave victims thinking that it affects only themselves so must be due to a personal failing. Example 3 was coded as negative.

Example 3 The equal opportunities watchdog has received more than 300 inquiries since mid-June about accusations against police of sexual harassment, but no complaints from actual VICTIMS, its chairman said yesterday. South China Morning Post (vLNSH500, VICTIM, 2019-10, 73)

The three examples above come from newspaper articles that provide context and source that assists with coding decisions. The situation becomes more complex when coding tweets because isolated from their thread they may lack key information that sets them into context. Where context is available tweets often seem primarily intended to attack others, often US political opponents, with the victim/survivor positioned as secondary. In Example 4 Lucy Flores<sup>57</sup> had publicly referred to inappropriate touching by Joe Biden, however tweets commonly suggest that she had accused him of sexual assault which she had specifically denied and corrected. This fallacy was repeated so often that the author of Example 4 may have come to believe it. The author is a Democrat Biden supporter using this tweet to defend him but also attack his opponent Donald Trump, a Republican. On the one hand the calling out of *a massive disservice to actual victims* could be seen as positive prosody of *actual victims*. On the other hand the marked phrase *actual victims* brings into being the unmarked victim, suggesting Lucy Flores is not an *actual victim* but a bogus victim. There is also an attempt to silence using the phrase *the world is done hearing* and the vengeful *know what sexual assault really is*. This tweet was coded as having negative prosody. Some of the most aggressive tweets in these datasets are attacking individuals who have made accusations against a writer’s favoured politician or celebrity.

---

<sup>57</sup> Lucy Flores is an American lawyer and worker for Joe Biden’s presidential campaign.

Example 4 @LucyFlores @JoeBiden You're doing a massive disservice to actual VICTIMS of sexual harassment. The world is done hearing from you, @lucyflores perhaps you should spend some time with @realDonaldTrump so you know what sexual assault really is. I'm sure he'd love to grab you by the p\*ssy!  
<https://t.co/NfSQFoKV07> (vLTSH500, VICTIM, 2019-04, 112)

Tweets where both accuser and accused remain anonymous are also common and Example 5, which appears to be a fragment of a thread, is fairly typical of these. MSNBC is a broadcasting company and @mmehenriquez is a human rights lawyer but it's not clear who is accusing who of what. A reader might align with the *genuine victim* and view the tweet as a positive calling out or equally they might be worried that if they speak out their disclosure would be smeared with the negative prosody of *this kind of bullshit*. Concordances like this that could be read as positive or negative were coded as neutral as there was insufficient context to make any other decision.

Example 5 @MSNBC @mmehenriquez This kind of bullshit marginalizes genuine VICTIMS of sexual harassment and rape .... (vLTSH500, VICTIM, 2019-04, 68)

A few concordances ("N/A" in Table 29 and Table 30) turned out to be unrelated to sexual violence, with the majority referring to the title of a TV series based loosely on a desert island shipwreck scenario. These either did not refer to sexual violence or were indecipherable like the one in Example 6. These concordances were coded as "N/A" and replaced by the next randomised concordance. Some authors appeared to be using the tag #MeToo in a way unrelated to the #MeToo movement.

Example 6 @sansbarrier Stalin and chinmayi most affected VICTIM in misa and #MeToo. (Twitter\_MT, VICTIM, 2019-10)

The next section discusses the initial analysis of +/- prosody of concordances.

### 4.2.3 Pilot of +/- prosody breakdown

Coding of the +/- prosody flags was carried out in batches of 100. The first batch are referred to as the vLPilot datasets<sup>58</sup>. The vLPilot datasets were used to fine-tune the process.

For all datasets there were more positive codes than negative and neutral combined, and Nexis datasets always had higher frequencies of positive codes than their Twitter equivalents (Table 29).

---

<sup>58</sup> The naming scheme for corpora and datasets is provided in Appendix B.

Dataset	KWIC	Positive	Neutral	Negative	N/A
vLTSHPilot	VICTIM	75	15	10	1
	SURVIVOR	84	8	8	5
vLTMTPilot	VICTIM	73	5	22	6
	SURVIVOR	81	7	12	3
vLNSHPilot	VICTIM	86	10	4	0
	SURVIVOR	100	0	0	3
vLNMT_Pilot	VICTIM	81	12	7	2
	SURVIVOR	99	1	0	24

Table 29, +/- prosody of concordances in the vLPilot datasets by VICTIM and SURVIVOR KWIC.

Additionally, SURVIVOR returned more positive concordances than VICTIM for all vLPilot datasets.

In the second batch there were significant increases in positive frequencies for SURVIVOR within vLTSH\_101-200 and vLTMT\_101-200 (Table 30). Although as with the first batch, the vL\_101-200 datasets also produced more positive codes than negative and neutral codes combined.

Dataset	KWIC	Positive	Neutral	Negative	N/A
vLTSH_101-200	VICTIM	83	2	15	2
	SURVIVOR	95	1	4	3
vLTMT_101-200	VICTIM	54	8	38	1
	SURVIVOR	91	1	8	1
vLNSH_101-200	VICTIM	96	2	2	0
	SURVIVOR	100	0	0	6
vLNMT_101-200	VICTIM	78	8	14	4
	SURVIVOR	90	10	0	0

Table 30, +/- prosody of concordances 101-200 from the datasets by VICTIM and SURVIVOR KWIC.

In contrast, vLTMT\_101-200 for VICTIM dropped to 54 positive codes in 101-200 from 73 in the vLPilot. Within vLNMT\_101-200 both VICTIM and SURVIVOR fell from 81 to 78 and 99 to 90 respectively. This positioned vLTMT\_101-200 SURVIVOR as slightly more positive than its Nexis equivalent counter to the general trend of Nexis corpora being the more positive data source.

#### 4.2.4 Third-party cross check of +/- prosody coding

In view of the variances between the +/- prosody codings of the vLPilot and vL\_101-200 datasets I requested that a colleague<sup>59</sup> carry out a checking exercise on vLNSHPilot SURVIVOR. He was provided with the study's definitions of positive, neutral, negative and N/A,

<sup>59</sup> Usman Maravia, a colleague who was concurrently doing a corpus linguistic PhD.

described in section 4.2.2. Interestingly, his coding was remarkably close to mine with only four different decisions.

I had coded Example 7 as N/A because it was a short headline referencing the web page “Survivors’ Stories” rather than people and it wasn’t clear whether the controversy cast survivors in a positive or negative light. My colleague coded it as positive.

Example 7 “SURVIVORS’ Stories sheds light on sexual assault amid controversy”  
The Leader-Post (Regina, Saskatchewan) (vLNSHPilot, SURVIVOR, 2020-12)

The TV show “Survivor Mexico” (Example 8) created some confusion as participants had been involved in a sexual harassment scandal. Both my colleague and myself were unsure, but we agreed that the accusation seemed peripheral to the *personalities*. Both feeling uncertain, confirmed my decision to exclude this concordance from the dataset.

Example 8 “The personalities who will be part of the reality show SURVIVOR México 2021 left for República Dominicana to begin filming, which continues despite the allegation of sexual harassment against TV Azteca producer Hernán “N”.”  
CE Noticias Financieras English (vLNSHPilot, SURVIVOR, 2021-03, N/A)

Title IX<sup>60</sup> is widely debated in the US, occurring 240 times in the NexisSH\_1Yr corpus, but not so well known elsewhere. Having followed the Title IX debate for two years I read Example 9 as positive as a calling out of the Devos amendments which create difficulty for survivors of campus sexual violence trying to report. My colleague, like many outside of the US, was unaware of the Title IX context and coded this as ‘N/A’.

Example 9 The revisions to Title IX DeVos and the Department of Education have made sends one clear message to sexual assault SURVIVORS: that they should not step forward.  
The Highlander: University of California - Riverside  
(vLNSHPilot, SURVIVOR, 2020-05, 15)

Bearing in mind the nature of the texts, I was satisfied with the level of agreement between our codings. My colleague and I are both UK-based and with similar educational backgrounds so it seemed reasonable that our initial responses to these concordances would be fairly similar. Our divergence around less mainstream topics is likely to reflect the range of interpretation by general readers.

The concordances that were coded as N/A are explored in the next section.

#### **4.2.5 Prosody breakdown of vL500 datasets**

The full breakdown for the vL500 datasets given in Table 31, shows that every SURVIVOR dataset has more positive representations than the VICTIM equivalent. Thus the answer to RQ1A is yes, SURVIVOR is more frequently represented as positive than VICTIM.

Table 31 also shows that vLTMT500 is the most negative dataset for both VICTIM and SURVIVOR. The highest frequencies of negative concordances is for vLTMT500 VICTIM (18%) which is

---

<sup>60</sup> Title IX is a piece of US equality legislation dating from the 1970s, and references to it in this study generally concern amendments made by US Department of Education Secretary Betsy DeVos in 2017 changing the standard of evidence required for reporting from “preponderance of evidence” to “clear and convincing”. For a crime that usually takes place in private, this presents many survivors with a problem.

more than three more than vLNMT500 VICTIM (5%). Additionally, the vLTMT500 VICTIM dataset is the only one where there are more negative than neutral codings.

Corpora	KWIC	+/- prosody coded concordances frequency (% of 500)		
		Pos	Neg	Neutral
vLTSH500	VICTIM	424 (85%)	30 (6%)	46 (9%)
	SURVIVOR	473 (95%)	11 (2%)	16 (3%)
vLTMT500	VICTIM	374 (75%)	90 (18%)	36 (7%)
	SURVIVOR	444 (89%)	27 (5%)	29 (6%)
vLNSH500	VICTIM	482 (96%)	5 (1%)	13 (3%)
	SURVIVOR	500 (100%)	0 (0%)	0 (0%)
vLNMT500	VICTIM	439 (88%)	25 (5%)	36 (7%)
	SURVIVOR	486 (97%)	1 (0%)	15 (3%)

Table 31, +/- prosody coding of vL500 datasets

Considering that #MeToo on Twitter was launched to be a safe space for victims to speak out it is disappointing that vLTMT500 contains the most negative codings for both VICTIM (18%) and SURVIVOR (5%).

#### 4.2.6 Excluded concordances

A close reading of the +/- prosody coding led to a number of concordances being excluded from analysis. I was initially concerned about the high (59) number of exclusions from the vLNSH500 SURVIVOR dataset (Table 32). Of the 59 exclusions 55 (93%) were found to be referring to the "Survivor Mexico" reality TV show. For the other dataset/KWIC combinations just over half of N/A codes referred to the TV show.

Van Leeuwen coded datasets	KWIC	Coded as "N/A"	TV show as % of "N/A"
vLTSH500	VICTIM	0	0%
	SURVIVOR	7 (5)	71%
vLTMT500	VICTIM	3 (0)	0%
	SURVIVOR	11 (6)	55%
vLNSH500	VICTIM	2 (0)	0%
	SURVIVOR	6 (4)	67%
vLNMT500	VICTIM	5 (0)	0%
	SURVIVOR	59 (55)	93%

Table 32, Excluded (N/A) concordances  
numbers in brackets relating to the "Survivor" TV show.

Within each dataset SURVIVOR always has more exclusions than VICTIM. None of the VICTIM texts contained references to the "Survivor Mexico" program suggesting that these were not discussing victimisation.



According to the van Leeuwen Social Actor Network discussed in section 3.4.1 Passivation is subdivided into the mutually exclusive Subjection and Beneficialisation. Subjection can manifest as Participation, Circumstantialisation or Possessivation, while Beneficialisation can also appear as Participation (2008, p. 34). Van Leeuwen refers to the complexity of language use in the wild, suggesting that discourses are not bound to a single fixed purpose by the author or subject to a single interpretation by the audience, both being influenced by the interlocutors' social context (2008, p. 124). While applying the van Leeuwen social actor codes within this study revisiting a coding often resulted in a shift of perspective as I will now try to illustrate using some of the more frequent co-occurring codes.

The most common code pairings involve Subjection, for instance Example 11 had originally been coded as Subjection because of the *empowerment and supporting* that SURVIVORS were subject to. However, on re-coding SURVIVORS were read as being Beneficiaries of Panhellenic's *reaching out to CARE* because the empowerment and support is conditional on the organisation CARE practising what it preaches.

Example 11 Kanumilli said Panhellenic has reached out to CARE to ensure that Panhellenic practices what it preaches when it comes to women's empowerment and supporting SURVIVORS.

The California Aggie: University of California  
(vLNSH500, SURVIVOR, 2020-10, 62)

Activation also appeared frequently as part of multi-coded concordances often paired with unlikely social actors. For example in the initial coding Example 12 was read as Subjection because of the *burden... that SURVIVORS shoulder*. On the re-coding the action of *carried a twin-sized mattress* was foregrounded and identified Activation as a voluntary act of defiance.

Example 12 To convey the burden of sexual assault and the feelings of violation that SURVIVORS shoulder, Sulkowicz carried a twin-sized dormitory mattress wherever she went for the remainder of her Columbia career, including across the Commencement stage.

The Columbia Spectator : Columbia University  
(vLNSH500, SURVIVOR, 2021-01, 71)

The judgement associated with Appraisal is often linked with passivating circumstances. In Example 13 the initial coding was Appraisal because SURVIVORS are evaluated as *justified*. However, on the second coding Beneficialisation seemed more prominent as *women have been systemically failed*. Indifferentiation was also coded because of the solidarity implied in *sharing their testimonies... as a form of self-defence*.

Example 13 Institutional mechanisms have systemically failed to protect women or provide justice, the verdict reasoned, and, therefore, SURVIVORS are justified in sharing their testimonies on media or social media platforms as a form of self-defence.

Indian Express (vLNSH500, SURVIVOR, 2021-02, 18)

A more complicated situation in Example 14 had initially been coded as Abstraction because the KWIC was part of a web page title rather than a person. However, on revisiting it this particular *woman* as one of the SURVIVORS and her story suggested Beneficialisation by the *Regina man...accused of sexual assaulting*. A third code of Subjection was also applied because he had *filed a lawsuit against* the author of her account.

Sexual violence and the gender pay gap  
 4 Prosody and sexual violence

Example 14 On Aug. 10, a Regina man who had been named and accused of sexual assaulting a woman on the SURVIVORS’ Stories page filed a lawsuit against the account’s authors and Facebook, which owns Instagram.

The Leader-Post (Regina, Saskatchewan) (vLNSH500, SURVIVOR, 2020-12, 25)

The slightly fluid nature of the interpretations led to the conclusion that the longer spent analysing a concordance the more possible social actor types that were likely to appear, and this would be inappropriate for tweets that are written and read in a few seconds as discussed in section 3.2.1. Hence the codings appearing in initial analyses were prioritised as valid as any.

Within the constraints of the current study it was not possible to analyse all 29 social actor types. In order to allow sufficient depth of analysis I narrowed my focus down to the social actor types that summed to more than 100 occurrences across the vL500 datasets for either VICTIM, SURVIVOR, or PAY GAP (Table 33). The six social actor types reaching this threshold are: Activation, Appraisalment, Beneficialisation, Indifferentiation, Subjection, and Suppression and these are described in section 3.4.1, and they provided 5,012 social actors representations.

Dataset	KWIC	Activation	Appraisalment	Beneficialisation	Indifferentiation	Subjection	Suppression
vLTPG500	PAY GAP	24	23	193	10	14	161
vLNPG500	PAY GAP	26	5	371	2	27	22
vLTSH500	VICTIM	53	53	129	24	192	0
	SURVIVOR	72	19	232	76	93	0
vLTMT500	VICTIM	63	112	156	21	123	0
	SURVIVOR	72	30	189	101	80	0
vLNSH500	VICTIM	101	5	271	6	130	0
	SURVIVOR	111	3	282	7	115	1
vLNMT500	VICTIM	101	82	189	7	152	0
	SURVIVOR	134	1	229	19	126	0
Total	PAY GAP	50	28	564	12	41	183
Total	VICTIM	318	252	745	58	597	0
Total	SURVIVOR	388	53	932	203	414	1

Table 33, Highest frequency social actor representations  
 Combines positive, neutral and negative representations. <sup>61</sup>

<sup>61</sup> Note that figures displayed in tables may not align with totals due to some concordances being coded to multiple values.

### 4.3.2 Third-party cross-check of van Leeuwen social actor types

A third party check was conducted for the van Leeuwen codes, similar to that described in section 4.2.4, with my colleague<sup>44</sup> assigning van Leeuwen social actor types to the vLNSHPilot SURVIVOR dataset. This showed that we had coded the same social actor types for 76 of the 100 concordances leaving 24 with different interpretations. Fourteen of the differences involved types of Passivation, 11 of which were alongside a non-Passivation actor type, and three also involved combinations of Activation plus Indetermination social actors. Nine of the 11 Passivation concordances were paired with Activation with the contradiction arising from the ambiguity of text which is discussed next.

Modal verbs appeared to drive several of the differences in interpretation, for instance Example 15 where *must be able to* can be read as Activation of a duty or an instruction to be enacted by the survivor. Read in the wider context of the following sentences however it suggests meaning akin to “must be allowed to” which passivates SURVIVORS.

Example 15 Minister for Women, Victoria Atkins, said: Sexual harassment is wrong and SURVIVORS must be able to share their stories. This survey will help us build a clear picture of who is affected and where. Working together with business, we can stamp it out.

Thai News Service (vLNSH500, SURVIVOR, 2020-01, 36)

The concordance in Example 16 was coded initially as Activation because of the verb *waiting*. The alternative interpretation had *wait* repositioning survivor as having been the beneficiary of investigations that had been delayed, resulting in a coding of Beneficialisation. This was strengthened by the *had to* phrase suggesting they were compelled with no choice.

Example 16 The letter also mentioned how Judicial Affairs is not representative of the diverse BU community, and SURVIVORS had to wait months for investigations to resolve.

The Daily Free Press: Boston University  
(vLNSH500, SURVIVOR, 2021-01, 102)

In Example 17 the personalised social actor type Indetermination uses *some survivors* which serves to anonymise. The *share their experience* was also interpreted as Indifferentiation because it carries the suggestion of support. The need for anonymity recognises those who have been victimised, but who are likely to regard speaking out publicly as being too risky.

Example 17 Some SURVIVORS fear retaliation or do not want to engage in a process, so being able to share their experience in an anonymous way can be a powerful step for them.

The Daily Barometer: Oregon State University  
(Nexis\_SH SURVIVOR, 2020-02, 95)

While closer alignment of coding between the initial and the third-party would have been ideal the type of differences suggest that some differences between coders or different rounds of coding with the same coder are almost inevitable and these will be explored further in section 4.3.2.1.

#### 4.3.2.1 Addressing subjectivity

The writers contributing news articles and tweets to this study and their audiences represent a wide range of political, geographic, educational, and economic perspectives, resulting in a

diversity of interpretations. Furthermore, the coding of concordances for this study suggests that a single reader's perspective might shift depending on what they've just read, seen on the news, or experienced on the train travelling home. For example a reader's interpretation of Example 9 discussing Title IX DeVos amendment might change radically if they had just seen a news item reporting a violent campus assault. Fairclough does not elevate the researchers interpretation over the general reader, but instead alludes to the need for transparency within the process.

*"The analyst is doing the same as the participant interpreter, but unlike the participant interpreter the analyst is concerned to explicate what she is doing."*  
(Fairclough, 2001, p. 139)

Similarly van Leeuwen (2008, p. 124) acknowledges the difficulty stating that discourses are not bound to a single fixed purpose by the author or subject to a single interpretation by the audience, both being potentially influenced by the interlocutors' social context, and FPDA encourages researchers to embrace "multiple marginalised voices" (Baxter, 2003, p. 72, 73).

In a world of fast-moving online news media and shortening attention spans news articles are generated quickly for instant consumption. Likewise readers may not reflect on the content at length possibly dipping in and out of the text rather than reading the whole piece or in the case of Twitter only reading a single disconnected tweet in isolation from a linked thread or news page. With no fixed universal meaning likely to be possible or desirable I take the position that my initial response to the text at the time of coding is as valid as any other reader on a social media platform. The next section describes the selection of social actor types to be taken forward for analysis.

### **4.3.3 Frequency of van Leeuwen Social Actor types**

The selection of the six van Leeuwen's social actor types for analysis is described in section 3.4.1, along with their frequencies. Reading across the Totals from Table 33, Beneficialisation is the most frequent social actor type for VICTIM (745) SURVIVOR (932) and PAY GAP (564). Next most frequent is Subjection with VICTIM (597) and SURVIVOR (414) and then Activation with VICTIM (318) and SURVIVOR (388). There is then a divergence with Appraisal VICTIM (252) and Indifferentiation SURVIVOR (203). Suppression is almost exclusive to the PAY GAP datasets (183) and heavily skewed towards the Twitter data.

Social actor types with total instances (from all vL500 datasets) below 100 are not shown in Table 33, and are excluded from the analysis.

The variation in frequency between the datasets suggests different patterns of use and this will now be addressed in the second half of this chapter where positive representations of social actor types are discussed in section 4.4, and negative representations in section 4.5. Suppression will be explored in chapter 5, as part of the pay gap discussion.

## **4.4 Positive representations of VICTIM and SURVIVOR concordances**

When positive prosody is mapped onto van Leeuwen social actors (Table 34) the picture for the positive vL500 datasets is similar to that seen for the full vL500 datasets (Table 33). This is because a high percentage of the concordances are coded as positive. Consequently Beneficialisation social actor types dominate the positive representations for both VICTIM and

SURVIVOR followed by Subjection and then Activation. Drilling down further SURVIVOR has higher frequencies of Beneficialisation, Activation and Indifferentiation than VICTIM while VICTIM has higher frequencies of Appraisalment and Subjection than SURVIVOR. This suggests that SURVIVOR is more closely associated with support, agency, and solidarity than VICTIM.

Dataset	KWIC	Activation	Appraisalment	Beneficialisation	Indifferentiation	Subjection
vLTSH500	VICTIM	50	39	124	23	186
	SURVIVOR	69	17	229	73	88
vLTMT500	VICTIM	52	57	146	20	102
	SURVIVOR	69	20	184	97	75
vLNSH500	VICTIM	98	5	270	6	128
	SURVIVOR	111	3	281	7	115
vLNMT500	VICTIM	89	27	179	6	131
	SURVIVOR	133	1	229	19	125
Totals	VICTIM	290	128	719	55	547
	SURVIVOR	382	41	923	196	403

Table 34, Highest frequency of positive representations  
Within the sexual violence vL500 datasets.

Note that there are multiple codings on some concordances, and low frequency social actor types were not analysed.

The sexual harassment social actor types: Appraisalment; Indifferentiation; and Passivation, map directly to themes. Additional sub-themes arose during analysis: Blaming/shaming; Consequences; Playing (victim); Speaking out; and Victim vs survivor.

The following sections explore positive social actor representations within VICTIM and SURVIVOR concordances of the vL500 datasets in order to address RQ1B. Negative representations are discussed in section 4.5.

#### 4.4.1 Activation

Activation social actor types are characterised by their association with dynamic roles within an activity (van Leeuwen, 2008, p. 33) making them particularly important for RQ1C in investigating the claim that the SURVIVOR label is associated with greater agency.

The data in Table 34 indicates that the highest frequencies of Activation occur in the two Nexis SURVIVOR datasets. Additionally SURVIVOR is represented more frequently than VICTIM across all datasets, represented 49% more frequently in vLNMT500, 38% more frequently in vLTSH500, 33% more frequently in vLTMT500, and within vLNSH500 13% more frequently. The frequencies of Activation seen in vLTSH and vLTMT are surprisingly close with SURVIVOR returning 69 instances in both and VICTIM differing by only two between vLTSH500 (50) and vLTMT500 (52).

The qualitative review of the Activation concordances found the dominant themes to be: Speaking out; the Consequences of speaking out; Playing victim; Blaming/shaming; and

distinctions between VICTIM and SURVIVOR. Multiple themes were applied to concordances where they applied. These main themes (Table 35) are discussed in the following sections.

Dataset	KWIC (positive only)	Activation social actors				
		Speaking out	Consequences	Play	Blaming/ Shaming	Victim vs Survivor
vLTSH500	VICTIM (32)	16	9	3	1	2
	SURVIVOR (69)	40	8	0	0	0
vLTMT500	VICTIM (52)	13	10	2	0	0
	SURVIVOR (69)	20	7	0	0	0
vLNNSH500	VICTIM (98)	56	22	0	1	2
	SURVIVOR (111)	54	14	0	0	3
vLNMT500	VICTIM (89)	34	15	0	2	0
	SURVIVOR (133)	58	35	0	0	6
Totals	VICTIM	119	56	8	4	3
	SURVIVOR	132	66	0	7	9

Table 35, Main themes for positive Activation representations within the vL500 datasets.

This data was gathered across 2019, coinciding with Tranchese’s (2023, pp. 380) final year data in which, relative to her earlier 2008-11 data, she found that *allege, tell, speak, say* and *describe* were either appearing, or showing increased use. She found that the main changes in her data occurred from 2011 as a result of changes in journalistic practices favouring the accused in the rising number of celebrity cases (Tranchese, 2023, pp. 164, 218, 400).

**4.4.1.1 Speaking out**

Speaking out is a broad theme identified by the use of:

- communication verbs such as *described, repeated, say/said, alleged;*
- phrases describing communication, for example *filed an affidavit, gave an impact statement;* and also
- phrases that imply communicating one’s story, including *opened up, coming forward, reached out, rise up, and fight back.*

Table 36 lists the top 10 Speaking out related verbs and verb phrases associated with VICTIM and SURVIVOR within the Twitter datasets, and Table 37 mirrors this for the Nexis datasets. Along with higher frequencies the Nexis datasets contain wider ranges of verbs expressing communication.

vLTSH500		vLTMT500	
VICTIM (of 15)	SURVIVOR (of 40)	VICTIM (of 13)	SURVIVOR (of 20)
<i>tell (4)</i>	<i>spoke (5)</i>	<i>testify (2)</i>	<i>come/step forward (4)</i>
<i>come/step forward (4)</i>	<i>talk/tell/told (4)</i>		<i>speak (2)</i>
<i>speak (2)</i>	<i>standing/stepping (3)</i>		<i>talk (2)</i>
	<i>use (3)</i>		<i>tell/told (2)</i>
	<i>voice (2)</i>		
	<i>signed (2)</i>		
	<i>push (2)</i>		
	<i>share (2)</i>		

Table 36, Comparing Twitter positive Activation verbs that occur more than once within the VL500 Twitter datasets.

The sequence of verbs with the same frequency is not significant.

vLN5H500		vLNMT500	
VICTIM (of 56)	SURVIVOR (of 54)	VICTIM (of 64)	SURVIVOR (of 58)
<i>report (6)</i>	<i>report (10)</i>	<i>said/tell (11)</i>	<i>come/step forward (6)</i>
<i>come/step forward (6)</i>	<i>speak (6)</i>	<i>accuse (2)</i>	<i>share (3)</i>
<i>said/tell (5)</i>	<i>said/tell (4)</i>	<i>used (2)</i>	<i>voice (3)</i>
<i>filed (4)</i>	<i>come/step forward (4)</i>	<i>share (2)</i>	<i>report (3)</i>
<i>write (3)</i>	<i>call (3)</i>		<i>said/tell (3)</i>
<i>complained (3)</i>	<i>demand (2)</i>		<i>speak (3)</i>
<i>voice</i>			<i>reveal (2)</i>

Table 37, Comparing Nexis positive Activation verbs that occur more than once within the VL500 Nexis datasets.

The sequence of verbs with the same frequency is not significant.

Frequencies of the Speaking Out theme are higher within the Nexis corpora and VICTIM is more frequent than SURVIVOR. This is reversed within vLTSH500 and vLTMT500 where SURVIVOR is more frequent. Within the Nexis datasets VICTIM and SURVIVOR have some words in common, such as: *Tell, Come/step forward, speak, report, voice, and share.*

The phrase *coming/stepping forward* occurs for both VICTIM and SURVIVOR in all of the sexual violence datasets except for VICTIM in vLTMT500 and vLNMT500. It appears in various contexts, for example: the military, politics, and the entertainment industries. These concordances use *come/step forward* in referring to initiation of some kind of legal action or other public calling to account. Alongside the acknowledgement of this positive step forward there is often a sense of weariness or discouragement, illustrated in Example 18 by the use of *disheartening* and the repetition of *different*. The tone is similar in the use of *finally* in Example 19.

Example 18 glad we live in an era of #MeToo where it's easier for SURVIVORS to come forward and expose their abusers, but it feels so disheartening at the same time when it's same scandals, different day, different channel I, different people. When does it stop? (vLTMT500, SURVIVOR, 2020-10, 3)

Example 19 We have the mass awareness of human trafficking now because of VICTIMS finally coming forward under the #metoo movement. Stealing that credit for Trump is theft of valor of a far uglier kind. Trump has no record on human trafficking. (vLTMT500, VICTIMS, 2021-01, 496)

This combining of two emotions suggests the complex and dynamic voices that Baxter describes (2003, p. 72).

#### **4.4.1.2 Consequences**

The consequences of speaking out about unacceptable behaviour are often negative for both VICTIM and SURVIVOR and within both vLTSH500 and vLTMT500 a greater variety of descriptors are associated with VICTIM than for SURVIVOR (Table 38).

Fear, shame and blame feature in the vLTSH500 VICTIM dataset along with verb phrases elaborating on the theme of fear, for example: *don't speak up for years; alleging harassment against a famous person far outweigh the benefits; and it took me some 20 odd years before I had the courage to tell anyone*. These consequences are represented as deterring the victimised or their advocates from speaking out. In contrast, vLTMT500 for VICTIM includes examples that directly counter-accuse those speaking out referencing for example: *personal gain, and get funded*.

vLTSH500		vLTMT500	
VICTIM (of 9)	SURVIVOR (of 8)	VICTIM (of 10)	SURVIVOR (of 7)
<i>fear, shame, and afraid of employers retaliating</i>	<i>reluctant to do (reporting)</i>	<i>expected to carry (burden)</i>	<i>Can Never Un-Know</i>
<i>change her public story</i>	<i>behave in a way counter to expectation (2)</i>	<i>never forgets</i>	<i>resented my own body</i>
<i>face scrutiny</i>	<i>turn hostile</i>	<i>relives the event every day</i>	<i>have to be perfect to be believed</i>
<i>don't speak up for years</i>	<i>feel that Linfield is a safe place to report</i>	<i>will shape harassment policy</i>	<i>no longer being believed</i>
<i>usually blamed</i>	<i>don't always come forward</i>	<i>become abusers</i>	<i>struggle to be heard (2)</i>
<i>alleging harassment against a famous person far outweigh the benefits</i>	<i>feels angry</i>	<i>personal gain</i>	<i>Survivors of sexual violence don't owe anyone our stories</i>
<i>it took me some 20 odd years before I had the courage to tell anyone reporting</i>		<i>become frauds</i>	
		<i>get funded</i>	
		<i>aren't believed</i>	
		<i>find the courage to speak out</i>	

Table 38, Consequences of speaking out in Twitter Activation concordances within positive coded vL500 Twitter datasets.

The writer of Example 20 offers some reasons why reporting of sexual violence is delayed including: *fear, shame, and afraid of employers retaliating*, and the catch-all *variety of reasons* alluding to numerous other possibilities. A Danish workplace study found that career damage and negative responses from the local community were consequences of voicing complaints (Nielsen et al., 2024, pp. 5-7). The second sentence warns against the common complaint that the victim is interrogated more rigorously than the perpetrator.

Example 20 VICTIMS of rape and sexual harassment don't always come out right away for a variety of reasons including fear, shame, and afraid of employers retaliating. Don't question the victims. Question the creeps like Callaway instead. (vLTSH500, VICTIM, 2021-02, 201)

#MeToo broke in India during the period of data gathering with complaints predominantly focused on Bollywood. Although not mentioning Bollywood explicitly in Example 21 the tone suggests one of those scandals. A comparison is set up between Hollywood where #MeToo is located but *idle* and another place. The implication is that if victims *face scrutiny* in Hollywood then the situation will be worse elsewhere. While *face* could be agentic and strong this concordance is also coded as Subjection because *scrutiny* was inflicted on the wronged.

Example 21 ...wouldn't seek justice. Any other resume with multiple incidents of sexual harassment would have been kept under wraps and even with a credible source he still got the job. The #MeToo movement is idle in Hollywood and even then VICTIMS face scrutiny... (vLTSH500, VICTIM, 2019-09, 367)

Many vLTSH500 tweets concern incidents on US university campuses. These usually refer to a SURVIVOR rather than a VICTIM. The vLTSH500 SURVIVOR verb phrases include:

- *reluctant to (report);*
- *behave in a way counter to expectation;*
- *feel that Linfield is a safe place to report; and*
- *don't always come forward.*

All of these refer to unseen barriers to reporting like Example 22 where the final sentence, a rhetorical question, anticipates the response that survivors do not feel Linfield is a safe place to report. The formal language of the committee's statement creates an impression of distance from the experience of *survivors*.

Example 22 'Lastly, and most important, the Board wants to remind the university community that any and all complaints about sexual harassment or assault should be promptly reported,' the committee said in its statement. Why, would any SURVIVORS feel that Linfield is a safe place to report? (vLTSH500, SURVIVOR, 2019-08, 439)

An unusual reference to anger appears in Example 23. This illustrates the confusion often seen between definitions of *harassment* and *assault*. *Uncomfortable* suggests a very minor inconvenience emphasised by the speech marks, and this seems to be at odds with the writer's understanding of *sexual harassment*. A contrast is set up between *dramatics and body language* with *Her words* suggesting an attempt to deceive. The writer self-identifies as a *survivor* who is angry at another<sup>62</sup> who has previously implied a claim of sexual harassment.

Example 23 Glad I'm not the only SURVIVOR who feels angry about this. Her words may be that she doesn't think it was sexual harassment, but her *dramatics and body language* are purposeful with an intent to conflate being 'uncomfortable' with sexual assault. (vLTSH500, SURVIVOR, 2019-04, 498)

The vLTMT500 VICTIM concordances in the third column of Table 38 carry a different tone. Several verb phrases: *become abusers, personal gain, become frauds* and *get funded* appear to tarnish the victim's reputation and/or suggest deviousness. The writer of Example 24 expresses a polarised view of the world that refers to a variety of social actors. The *goes hand and hand* implies that *#MeToo and #lgbtq* activists conspire to *accuse the blacks, play victim and get funded*. Sexual harassment has historically had a strong racial element and the concept of MeToo was initiated by Tarana Burke, a black woman. Additionally, many voices on Twitter represent black and ethnic women as suffering most from sexual harassment and receiving least justice. The noun phrase *money grab game play* suggests that this is a well-known strategy whereby bogus accusers can *play victim and get funded*. In the final

---

<sup>62</sup> From the timing, the person referred to may be Lucy Flores, refer to Example 4.

sentence *I'm not a black celebrity so...* represents those with the power to influence as turning a blind eye.

Example 24 The #MeToo and the #lgbtq goes hand and hand. It's a money grab game play. Accuse the blacks, play VICTIM and get funded. I'm not a black celebrity so I can actually call it what it is. (vLTMT500, VICTIM, 2019-11, 495)

Other vLTMT500 verb phrases such as, *expected to carry*, *never forgets*, and *aren't believed*, speak of ongoing psychological consequences. Example 25 combines the sense of burden implied by *already been through so much*, with the suggestion of healing that would provide *the courage to speak*. However, the final three sentences emphasise the vulnerability of the *real people* harmed.

Example 25 Yes. I totally agree. Sexual assault VICTIMS have already been through so much before they find the courage to speak out. Harassment&bullying just adds to the pain. They aren't just words typed on a screen . We are real people. We have feelings. Words hurt #MeToo (vLTMT500, VICTIM, 2020-08, 422)

Staying with the vLTMT500 dataset but turning to SURVIVOR concordances these predominantly express a state of being, with belief and sharing of information featuring heavily. The verb phrases in the fourth column of Table 38 suggest an unhappy impotence. For example, *resented my own body* makes the survivor both object and subject of the verb phrase, and suggesting self-blame or shame. *Struggle to be heard* on the other hand emphasises the great effort expended, but which is still not enough to be fully *heard* and suggesting a form of silencing.

The full tweet around the final SURVIVOR verb phrase from Table 38 is given in Example 26. The dramatic *rip our pasts open* suggests a violent and courageous act, but also alludes to the pressure purportedly from West<sup>63</sup> to share details publicly. There is an interesting contrast between the phrase *don't owe*, implying an absence of obligation with the opening *I wish Women didn't have to* which suggests a reluctant admission that in reality women do *have to*. The final phrase *I'm telling mine* is a first person agentic statement suggesting the writer is taking control.

Example 26 I wish Women didn't have to rip our pasts open and show you everything and let you ogle our pain for you to believe us. - Lindy West, SURVIVORS of sexual violence don't owe anyone our stories. Here's why I'm telling mine. #MeToo? (vLTMT500, SURVIVOR, 2021-01, 308)

The discussion now turns to the Nexis corpora, and the consequences stated in those concordances are summarised in Table 39.

---

<sup>63</sup> Lindy West is a feminist author/comedian in the news in 2019 following the release of her book "The Witches Are Coming".

Sexual violence and the gender pay gap  
 4 Prosody and sexual violence

NexisSH		NexisMT	
VICTIM (of 22)	SURVIVOR (of 8)	VICTIM (of 16)	SURVIVOR (of 19)
<i>fear (6)</i>	<i>rid myself (2)</i>	<i>fear (4)</i>	<i>begin to heal (2)</i>
<i>career consequences (5)</i>	<i>burdened</i>	<i>bear the brunt</i>	<i>confrontations (2)</i>
<i>get used to it (2)</i>	<i>could not step forward</i>	<i>hide</i>	<i>face prejudice</i>
<i>driven to depression</i>	<i>systemic hurdles</i>	<i>not able to work</i>	<i>been threatened</i>
<i>mental health impacted</i>	<i>re-tell their stories</i>	<i>stay silent</i>	<i>painful experiences</i>
<i>difficulty getting justice</i>	<i>bias and discrimination</i>	<i>face consequences</i>	<i>took it for the team</i>
<i>experience retaliation</i>	<i>face obstacles</i>	<i>keep their secret</i>	<i>blame ourselves</i>
<i>experience retribution</i>		<i>repressed pain</i>	<i>avoid damage</i>
		<i>married</i>	<i>cut and bleed</i>
			<i>felt a tragic loyalty</i>

Table 39, Consequences of speaking out in Nexis Activation concordances

All 22 of the NexisSH VICTIM concordances in the first column of Table 39 are work place related. *Fear of speaking out, fear the consequences, and afraid to talk* verb phrases all referencing work place cultures and five specifically mentioning possible retaliatory consequences for their careers. Example 27 discusses the murder of Vanessa Guillen<sup>64</sup>. The issue being raised is that the reporting route for complaints is through a chain of command that includes the perpetrator who will have a vested interest in minimising this and future complaints. *Retaliation* suggests a response outside the official regulated procedure either by creating barriers to career progression and/or turning a blind eye to further abuse. Career damaging retaliation was found to result from calling out sexual violence in Danish workplaces (Nielsen et al., 2024).

Example 27 The former victim advocate and other women said the low assault numbers on the base are likely attributable to women believing that they cannot file a report, out of fear that it would be handled internally and result in retaliation against the VICTIM , not punishment for the perpetrator.

Philadelphia Daily News (vLNSH500, VICTIM 2020-07, 191)

In contrast to the VICTIM concordances all of the NexisSH SURVIVOR concordances discuss an anonymous generality of survivors and the culture surrounding them. The muting or silencing of individuals who speak out goes beyond preventing access to justice. Example 28 refers to the long-term toll taken by carrying the weight of that trauma. The writer describes the effort of maintaining concealment of the secret, emphasised by the repetition of *that burdens you*.

<sup>64</sup> Vanessa Guillen was a 20 year old US soldier, murdered at Fort Hood in 2020 by her supervisor who had been sexually harassing her.

Example 28 Xxxxx Xxxxx<sup>65</sup>, a survivor of childhood molestation, rape and domestic violence, says #MeToo has cleared the way for countless SURVIVORS like herself to speak openly about that trauma. “It’s the secrets that burden you. It’s the hiding that burdens you,” she said.

USA TODAY (vLN500, SURVIVOR, 2019-09, 278)

The dominant themes within the NexisMT VICTIM concordances in the third column of Table 39 are similar to the previous datasets, with approximately half referring to career consequences. *Fear, bear the brunt, not able to work, and repressed pain* all touch on the negative consequences that victims anticipate if they speak out. In Example 29 Tanushree Dutta<sup>66</sup> refers to the double-edged consequences Bollywood actresses face if they speak out; either being overwhelmed by *trauma* resulting from the media response or social and economic *ostracisation*. Unlike explicit discrimination ostracisation is likely to be very difficult to prove and address either legally or informally. While Tanushree is likely to be serious her rhetorical question implies she does not really believe *Bollywood has sleepless nights*.

In British English *smirk* would suggest an unpleasant smugness that feels a little awkward and unlikely here, but this might not be the same for other variations such as Indian English. As a brief deviation within the enTenTen21 the top modifiers of SMIRK are: evilly (9.5), smugly (7.7), inwardly (7.5), mischievously (7.4), wickedly (7.3), devilishly (7.2), triumphantly (7.0), slyly (6.5), arrogantly (6.5), and playfully (6.4) and predominantly from the fiction genre. As an actress working in fictional films in Bollywood, Tanushree Dutta may be more familiar with the language of the movie fiction genre, and that may be reflected in her everyday speech in this quote.

Example 29 Tanushree Dutta, who was the reason for #MeToo storm in India, gave an interview to Mid-Day where she counter-questioned the actor. “How come nobody in Bollywood has sleepless nights when a woman becomes a VICTIM of abuse and is not able to work due to the trauma and ostracisation,” stated Tanushree with a smirk.

Daily News & Analysis (vLNMT500, VICTIM, 2019-09, 237)

In Table 39 column 4 NexisMT SURVIVOR concordances cover an equal mix of workplace, Hollywood/Bollywood, high profile court cases, and general discussion about sexual harassment. There is a theme of self-blame and having difficulty recognising abuse evidenced in *blame ourselves, tragic loyalty, and avoid damage* (to others). The phrase *took it for the team* describes a victim taking 22 years of abuse because they didn’t want to let *the party* down. The picture painted in Example 30 is closely related to Example 29, and all the more chilling for its neutral tone. The writer suggests that survivors are either discouraged from reporting sexual harassment or anticipate a negative response. The use of *allows*

---

<sup>65</sup> The subject Xxxxx Xxxxx didn’t appear in a Google search so may not be a public figure, and hence their name has been omitted.

<sup>66</sup> Actress and former Miss India, who made an allegation against another actor early in the #MeToo movement in India.

suggests a culture that has naturalised sexual harassment resulting in survivors either remaining silent or leaving while perpetrators remain un-checked.

Example 30 Instances of sexual harassment, and of sexual assault, happen in our party. They don't often come to light, usually because the SURVIVOR leaves the party, but also because those survivors who stay have no official means of reporting such behaviour. "This allows abusers to continue in the party.

The Dominion Post (vLNMT500, SURVIVOR, 2019-09, 440)

#### **4.4.1.3 Playing victim**

Accusations of *playing the victim* are only associated with Twitter VICTIM concordances, there is no SURVIVOR equivalent. These were primarily coded as Activation because of the verb *play*, however most are also coded as Appraisal because the phrase is evaluative. The implication is that not only has the individual brought this upon themselves, but they acted with intent to either gain personal benefit, or to deliberately damage the accused, as previously seen in Example 24. Within these positive representations the phrase is used by advocates calling out this form of victim-blaming. However, taken from vLTSH500, Example 31 shows the reverse of the usual situation, where someone accused of predatory behaviour is directly represented as *playing victim* for failing to take responsibility for their actions and trying to position themselves as the wronged party.

Example 31 asking random people to date you/participate in sexual affairs with you is sexual harassment, them being minors just makes it worse. there is no "it's just a joke" or "y'all take everything too seriously" or playing VICTIM because you're facing backlash from your little "joke" (vLTSH500, victim, 2019-08, 42)

Claiming that sexual harassment is *a joke* adds insult to injury, firstly by blaming the victim for getting upset and, and secondly for thinking badly of the perpetrator who has positioned themselves as having innocent intentions.

#### **4.4.1.4 Blaming/shaming**

*Victim-blaming* and *shaming* refers to longstanding rape myths which hold that any unwanted attention is the victim's fault arising from their choices, including their: clothes; location; and choice of companions. In SketchEngine the WordSketch algorithm identifies *blaming* and *shaming* as verbs enacted by *victims*. However, on viewing the concordances manually *victim-blaming/shaming* as a verb phrase always has the victim as subject, i.e. victim is being *blamed/shamed*. This is always in the context of a victim or their advocate calling out the victim-blaming behaviour, and hence this is positive semantic prosody. Although initially identified as Activation because VICTIM appeared to be the verb's subject, they remained as Activation because the victim is credited with having caused the attack by some error or omission in their behaviour.

In Example 32 an actress is taken to task for attributing *the behaviour of men to the way women dressed* resulting in their victimisation. Most of the *blaming* and *shaming* concordances are also assigned the Appraisal social actor type.

Example 32 Her remarks got her into a certain amount of trouble with other women in Hollywood who claimed she was ‘VICTIM-BLAMING’ - specifically, that she was blaming the way women dressed for the behaviour of men.

The Mail On Sunday (vLNMT500, VICTIM, 2019-09, 398)

Rape myths are more strongly rooted in some cultures than others. Example 33 refers to a news story from 2016 at the Mormon Brigham Young University which, by the standards of the rest of the USA and the UK, has a very conservative dress code and standards of behaviour (<https://policy.byu.edu/view/dress-and-grooming-standards>). Here the victim appears to have internalised the *victim-blaming* myth, and although stating that she has overcome *self-blame*, the statement that she *preferred not to be named* suggests that some self-blame or shame remains.

Example 33 One VICTIM, who preferred not to be named, shared with the Daily Universe her story of overcoming that sense of self-blame after she was sexually assaulted right before starting school at BYU.

The Daily Universe: Brigham Young University (vLNMT500, VICTIM, 2020-07, 82)

The single instance within vLTSH500, Example 34, depicts a particularly emotionally raw expression of self-blaming. The victim not only blames herself for the harassment, but also for being upset by it, and finally for expecting a supportive response from the people she told. The repetitive use of *fucking* here creates a sense of anger and hurt. The victim claims total agency for the situation, but it isn't clear what choices she could have made differently to change her outcome.

Example 34 (sexual harassment, VICTIM-BLAMING & self blame tw) i feel so fucking stupid for not being over what happened, and i feel so fucking stupid for letting it happen in the first place, and i feel SO fucking stupid for thinking amtrak would ever believe me or give a shit. (vLTSH500, VICTIM, 2020-03, 452)

Neither *blaming* nor *shaming* collocate with SURVIVOR.

#### **4.4.1.5 Victim vs survivor**

Concordances in this category highlight the distinction between the labels VICTIM and SURVIVOR, usually with a marked preference for one or the other. Table 35 showed that the vLNMT500 SURVIVOR dataset returned six VICTIM vs SURVIVOR examples, vLNSH500 VICTIM and vLTSH500 VICTIM both returned two examples, and the vLNSH500 SURVIVOR dataset contained three. Example 35 is interesting in that the writer offers explicit definitions suggesting that real world meaning may be evolving from traditional OED definitions (Oxford English Dictionary [electronic resource], n.d.-], n.d.-]).

Example 35 The word “survivor” evokes more positive emotions and coping mechanisms, she said. “ ‘VICTIM’ evokes pity; ‘survivor’ evokes admiration.”

Pittsburgh Post-Gazette (vLNMT500, VICTIM, 2019-09, 372)

While Example 35 suggests a simplistic choice between an unflattering option and an empowering alternative, Example 36 alludes to some fluidity between the two. This tweet describes a journey from being a victim, and *not being strong enough*, to *today when ALL victims* are survivors. However, the strength and agency of SURVIVOR comes at a price

suggested by the *should speak* which carries a tone of obligation or expectation to provide a narrative of overcoming victimisation.

Example 36 I had to deal with the consequence of not being strong enough and to fight for myself for months but today, ALL VICTIMS of sexual harassment should speak on their stories because baby you are a survivor.

(vLTSH500, VICTIM, 2020-06, 130)

The contrary position is given in Example 37 which uses the combined *victims/survivors* implying an equivalence. Here a relaxed and friendly tone is created with *y'all*, and the point being made is that *victims/survivors* are under no obligation to share their stories. The *if they decide on it* phrase credits individuals with agency without obligation.

Example 37 y'all know people can deal with and protect others from problematic individuals in other ways than public social media callout posts? y'all know VICTIMS/survivors of sexual harassment can choose to not share their stories if they decide on it?

(vLTSH500, VICTIM, 2020-01, 495)

These examples suggest that the distinction between VICTIM and SURVIVOR is important to those individuals themselves and also to others, but it seems the definitions may remain open to interpretation.

#### 4.4.2 Appraisalment

Referring back to Table 34, there are more than four times as many Appraisalment concordances within the Twitter vL500 (133) datasets as there are in the Nexis (36) equivalents, and 75% are associated with VICTIM concordances. vLTMT500 has 57 concordances and the collocates of VICTIM include: *real, actual, true, fake, deserves, pretend, right kind of, freak show* and *so-called*. These positive concordances use the adjectives in two ways, to either: call out someone attacking the specific victim being discussed like the individual positioned as *a freak show* in Example 38; or to accuse this victim of harming other anonymous and deserving *real victims* in Example 39.

Example 38 unusually contains a double appraisalment. In an earlier tweet the victim has been attacked as being a *freak show*, and here the attacker has in turn been labelled *disgusting*. *Freak show* seems to suggest that by speaking out, the victim has put themselves on show to the public for gain and in an unedifying way, which would be damaging to their reputation, or delegitimising (Oxford English Dictionary [electronic resource], n.d.-o).

Example 38 Referring to people raising concerns about being potential #metoo VICTIMS as a "freak show" is disgusting. (vLTMT500, VICTIM 2020-10, 141)

Twelve of the 57 vLTMT500 concordances contain the expression *real victim* as in Example 39, which suggests that *she* has broken some unspecified rule of victim behaviour. Furthermore, in falsely claiming victimhood *she* has damaged other unspecified but deserving victims of GBV (gender based violence). The dramatic *Oh My Gawd* indicates that the writer could be bored, irritated, or shocked (Collins Dictionary, n.d.-a), and this is amplified by the repeated punctuation marks (Crystal, 2006, pp. 37, 38, 95, 170). Readers might be bored or irritated because they believe that bogus victims speak out too much, as in Example 4 and Example 161, or they might be shocked that bona fide victims have to speak out so often.

Example 39 I said that as well...she alone has done soooo much damage to the #MeToo movement????????? & to the real VICTIMS of GBV!!!! Oh My Gawd!!!!???.

(vLTMT500, VICTIM 2020-02, 401)

Crystal suggests that using punctuation like this supports prosody via a limited range of emotions such as surprise, puzzlement, and “attitude”, it is not clear which of those is being expressed here. In contrast, a study based on youth culture found the repetition of ! and ? punctuation marks may convey extreme wonder/dissatisfaction and extreme indignation respectively (Bilenka & Kolomiiets, 2021, p. 94). However, Vandergriff (2013, p. 11 (in pdf reader)) found her second language participants used multiple exclamation marks to express assertiveness or excitability. An alternative view from Angouri & Tseliga (2010, p. 70) is that multiple punctuation may be a creative compensation for the absence of gesture and tone in an online environment.

In a later exploration, a search of the TwitterMT corpus using the term *real victim* returned 26 occurrences, all discrediting or denying a victim. A significant number of these attack female celebrities who have raised legal cases such as Amber Heard and Christine Blasey-Ford. In contrast, there are only five instances of *real survivor*, four of which are first person sentences using the phrase “speaking as a real survivor...”. Use of the adjective *real* is discussed further in section 4.5.2.

#### 4.4.3 Indifferentiation

Across all positive datasets, there are more Indifferentiation representations for SURVIVORS than for VICTIMS, with more than three times as many within the Twitter vL500 datasets, and only a slightly lower ratio in the Nexis vL500. Indifferentiation is found in 97 concordances within vLTMT500 and in 73 in vLTSH500, however there are only six each in vLN500 and vLNMT500. The hallmark of Indifferentiation is the sense of solidarity created, and in vLTMT500 this often manifests through the use of personal pronouns such as ‘we’ and ‘our’, suggesting identification with another, or a common bond. Indifferentiation is a theme in it’s own right.

In Example 40 *we* is used four times in 14 words as the writer speaks as a survivor, and for other survivors. The *loud & proud* suggests that the writer has come forward with their story publicly. The writer follows this up with *jst gotten started* which suggests that they have more to say, and intends to carry on doing so.

Example 40 ... We’re here & we’re not going anywhere. We’re loud & proud. We are SURVIVORS. & we’ve jst gotten started. (vLTMT500, SURVIVOR, 2020-07, 184)

Bouvier (2022, pp. 185, 186) asks who *we* are, in this case perhaps: survivors, feminists, women, or a specific group. She explained the importance of this in creating “the polarized ‘our people’”.

In Example 41 the writer introduces themselves with *I am a #MeToo survivor*, establishing the personal experience from which to speak for others. In the context of the opening sentence’s bad smell metaphor indicating suspicion *We don’t change our stories* suggests that Reade has changed her story, and this doesn’t fit the writer’s view of the world. This tweet suggests that survivors all behave in the same way, and those taking a contrary position are not

believable. It is possible that these two appraisements in the third and fourth sentences believing *Ford* (accusing Kavanaugh, a Republican nominee) but disbelieving *Reade* (accusing Democrat ex-President Biden) may reflect political motivation. Palomino Manjón's (2024, p. 120) study based on the Blasey-Ford vs Kavanaugh case found that support was politically polarised, with Republicans viewing #MeToo as a political weapon used by Democrats.

Example 41 This stinks like a week old decaying body of a skunk left in hot weather. I am a #MeToo SURVIVOR. We don't change our stories. I believe Ford. I have trouble believing Reade. (vLTMT500, SURVIVOR, 2020-05, 320)

Some writers claiming to speak for others may have relevant experience to justify this. The author of Example 42 sets out their allegiance with several groups to represent herself as a *strong survivor who fights for victims*. Here, *victims* are represented as needing someone to defend them, while the agentic SURVIVOR is altruistically providing that service. The opening *Sorry, sir* warns that the writer is about to take a contrary position, which is then announced by *But*. The accusation discussed here is *without proof, by no one involved*. Here, there is a reversal with the accused repositioned as the "real" victim (Shoaf et al., 2024, p. 1910).

Example 42 Sorry, sir, but you are full of it! I am a feminist, and a SURVIVOR of sexual harassment, sexual abuse, physical abuse, etc.! I am a proud and strong woman who fights for victims. But the only "victim" I see here is the one accused, without proof, by no one involved! #metoo (vLTSH500, SURVIVOR, 2019-06, 387)

#### 4.4.4 Passivations

The Passivations Beneficialisation and Subjection are discussed together in this section as these concordances are often either present as existing on a continuum with a significant grey area, or as part of an "on the one hand, or on the other hand" structure. Referring back to Table 34, Passivations appear with much higher frequencies than any other social actor type within both VICTIM and SURVIVOR concordances. Subjection always appears at a higher frequency for VICTIM than SURVIVOR, and is always weighted towards the Nexis and #MeToo vL500 datasets. Beneficialisation ranges in frequency from SURVIVOR having almost three times that of VICTIM within vLTSH500, to only 4% more within vLNSH500, refer to Table 40. In contrast, Subjection is more frequent for VICTIM than SURVIVOR across all datasets, ranging from 72% for vLTSH500 down to 5% for vLNMT500.

Dataset	KWIC	Beneficialisation	SURVIVOR / VICTIM	Subjection	SURVIVOR / VICTIM
vLTSH500	VICTIM	82		151	
	SURVIVOR	229	179%	88	72%
vLTM500	VICTIM	146		102	
	SURVIVOR	184	26%	75	36%
vLN500	VICTIM	270		128	
	SURVIVOR	281	4%	115	11%
vLNMT500	VICTIM	179		131	
	SURVIVOR	229	28%	125	5%

Table 40, Distribution of positive Passivations between VICTIM and SURVIVOR within the positive vL500 datasets.

A review of the Beneficialisation concordances within the vLTSH500 dataset suggests they are positive and supportive for both VICTIM and SURVIVOR. Here the SURVIVOR concordances have a kinder and often more personal tone. In contrast, the VICTIM counterparts can appear somewhat adversarial in their defense of victims, see Example 43 and Example 44. In the former SURVIVORS benefit from the writer *holding space forever and always*, and in the latter they benefit from the writer advocating for their need to be *heard in court & get a Closure Verdict!!*.

Example 43 Holding space forever and always for SURVIVORS of sexual harassment and assault. (vLTSH500, SURVIVOR, 2021-02, 15)

Example 44 Bill Cosby is in Prison; Trump is not above the Law, these VICTIMS need to be heard in court & get a Closure Verdict!! (vLTSH500, VICTIM, 2019-06, 13)

Within the vLTSH500 dataset's Subjection concordances, VICTIM is generally associated with aggressive verbs and verb phrases such as: *bad mouthed, slandered, bullying, and slap* see Example 45. In contrast, SURVIVOR concordances generally employ a less aggressive, though still unpleasant, tone including: *asked, demeaned, outed, face, silenced, bear the cost, and humiliated*. Example 46 comes across as sarcastic and patronising with *I'll pray for you to have more compassion & empathy*, which is an underhand way of saying you lack these attributes.

Example 45 Awarding the Medal of Freedom to #GymJoran is a slap in the face to all VICTIMS of sexual assault. (vLTSH500, 2021-01, VICTIM, 16)

Example 46 And I'll pray for you to have more compassion & empathy, esp. for SURVIVORS of sexual harassment & assault. (vLTSH500, SURVIVOR, 2020-02, 53)

The next section switches focus to the concordances coded as negative.

**4.5 Negative representations of VICTIM and SURVIVOR concordances**

A consequence of the selection criteria, described in section 4.3.3, is that +/- prosody codes for some social actor types are predominantly positive, with fewer or no negative examples for some datasets (Table 41). Generally, there are more negative social actor types for VICTIM than SURVIVOR, with the exception of Indifferentiation which has very low frequencies and is driven by the Twitter SURVIVOR datasets.

Dataset	KWIC	Activation	Appraisal	Beneficialisation	Indifferentiation	Subjection
vLTSH500	VICTIM	1	12	3	0	4
	SURVIVOR	3	2	3	3	5
vLTMT500	VICTIM	11	55	10	0	21
	SURVIVOR	2	10	5	4	5
vLN500	VICTIM	3	0	1	0	2
	SURVIVOR	0	0	0	0	0
vLNMT500	VICTIM	9	3	6	0	9
	SURVIVOR	0	0	0	0	1
Totals	VICTIM	25	72	22	2	38
	SURVIVOR	6	12	8	7	11

*Table 41, Frequency of negative representations of social actor types within the sexual violence vL500 datasets. Note that there are multiple codings on some concordances, and low frequency social actor types were not analysed.*

Bearing in mind that the #MeToo hashtag was originated as a safe space for women to speak out, it may seem contradictory that 123 of the tweets in vLTMT500 coded as negative, more than any other dataset. However, #hashtags are accessible to everyone and challenges may occur and at times overwhelm the original purpose, contesting which social actors were foregrounded within the debate (Jones et al., 2022, pp. 6, 8, 9)

In order to address RQ1B, the following sections explore the negative representations of VICTIM and SURVIVOR for the social actors appearing in Table 41.

**4.5.1 Activation**

Negatively coded Activation concordances are more common at the intersection of Twitter, #MeToo and VICTIM, consequently vLTMT500 VICTIM has almost half of all of the negative Activation codes in Table 41. The predominant theme is Speaking Out, as it was for positive concordances, with the Consequences and Playing Victim still present albeit at lower frequencies (Table 42). Because of the low numbers, all of the verb phrases for these themes are given. In these negative examples there are no examples of blaming/shaming, or Victim vs survivor.

Dataset	KWIC	Activation		
		Speaking out	Consequences	Playing Victim
vLTSH500	VICTIM	0	0	<i>playing the victim game pretend to be victims</i>
	SURVIVOR	<i>demand claims taking credit</i>	0	0
vLTMT500	VICTIM	<i>she called the shots woman can say dear to a man (3) made an allegation</i>	0	<i>-she played victim a platform to play the victim</i>
	SURVIVOR	0	0	0
vLN500	VICTIM	<i>submitted a complaint revealed filed the complaint</i>	0	0
	SURVIVOR	0	0	0
vLNMT500	VICTIM	<i>now to say... accusing others claims speak out</i>	<i>become heroes</i>	<i>playing the victim</i>
	SURVIVOR	0	0	0
Totals	VICTIM	12	1	5
	SURVIVOR	3	0	0

Table 42, Dominant themes within the negative Activation concordances within vL500 datasets.

Note: some concordances contain more than one theme.

The following sections discuss the concordances referred to in Table 42.

#### 4.5.1.1 **Speaking out**

Speaking out occurred in the 3<sup>rd</sup> person in all but two concordances, and when associated with VICTIM it was expressed as the speech verbs: *complaint*, *revealed*, *allegation*, *accusing*, *saying*, and *speak out*. Within SURVIVOR concordances, speaking out took the form of *demand*, and *claims*, and both instances were within the vLTSH500 dataset. The two concordances in Example 47 and Example 48, appear to be sarcastic or possibly attempting humour at the survivor's expense. Both position survivors as making frivolous and fraudulent claims to gain attention or financial benefits. Section 2.5.2 describes the serious deterrents to reporting faced by victims in the UK.

Example 47 Im a SURVIVOR of sexual harassment & I demand @twitter verify me & give me a free year of #TwitterPro?. You can send me #dogecoin to thank me for my bravery in coming forward at: ... (vLTSH500, SURVIVOR, 2021-02, 49)

Example 48 Oh, you mean like the sexual harassment “SURVIVORS” that never make a report to police and instead they make their baseless claims at social media? (vLTSH500, SURVIVOR, 2020-02, 188)

The South Korean quote in Example 49 was coded as negative because although it highlights a lack of justice, it appears to fall short of calling-out sexual harassment, instead it strikes a tone of description and acceptance.

Example 49 According to a statement released by a committee to deal with Yang’s alleged sexual harassment, a VICTIM submitted a complaint to the Seoul foundation and asked for punishment of Yang, an apology from the leader of the foundation and further measures to prevent similar incidents in the future. ... the foundation didn’t take any steps to actually punish Yang and closed the case claiming that it has no authority to investigate Yang because they are no longer under contract with one another. Korea Times (vLNSH500, VICTIM, 2020-06, 15)

Ocasio-Cortez, a US Democrat politician, is firstly described as a self-identified victim, referring to her agency in speaking out for herself, but then accused of being *similarly quiet* (similar to a list of US, female, and left-wing politicians) for failing to speak out on an unnamed situation in Example 50. The implication is that, contrary to expectations, Ocasio-Cortez puts her political career before her principals. According to a now deleted page, she failed to meet an expectation of speaking-out for a victim of Andrew Cuomo, the Democrat Governor of New York. The obligation to speak out for others, generally associated with SURVIVOR, is discussed further in section 7.4.

Example 50 Meanwhile, Alexandria Ocasio-Cortez - who recently revealed she had been a VICTIM of sexual assault - has been similarly quiet. MailOnline (vLNSH500, VICTIM, 2021-02, 20)

The defamation case against Kuta, in Example 51, was upheld by a US court. The expression *jump on the bandwagon* suggests that Kuta was trying to “to join in what seems likely to be a successful enterprise” (Oxford English Dictionary [electronic resource], n.d.-c). The final phrase, *get instant victim status*, suggests that making unfounded allegations is an easy and popular way to obtain some benefit. The reference to a *powerful man*, likely to have resources, further highlights the possibility of a material gain. No charges were ever filed against Wynn.

Example 51 A Woman Who Accused Former Casino Mogul Steve Wynn Of Rape Was Found To Have Defamed Him <https://t.co/MGg0LMGX3C> Kuta was also able to jump on the #MeToo bandwagon in 2018, getting instant VICTIM status simply because she made an allegation against a powerful man. (vLTMT500, VICTIM, 2020-03, 165)

Before deduplication the raw Twitter data contained many texts posted by Johnny Depp and Amber Heard fans, who were polarised to a simplistic level of support for their favoured celebrity. Although divorced at the time of this study, accusations of domestic abuse and

defamation remained in the media. Here, in Example 52, a Depp fan represents Heard as using her *white privilege to take advantage* and bring about a misdemeanour, in contrast to Depp (also white and privileged) who uses his *celebrity status for good*. Moreover, Heard is also portrayed as encouraging her friends to not just *lie*, but preceded by the redundant *falsely* to emphasise the point. In contrast to the initial representation of Heard as agentically manipulative and dangerous, the final sentence repositions her as *just SAD&pathetic*.

Example 52 @froggy57 Didn't cry over split milk.. Many of #JohnnyDepp fans know he's a hard worker NOT spoiled,used his celebrity status for Good ..While #AmberHeard used her white privilege to take advantage of #MeToo to play VICTIM &had her celebrity friends to falsely LIE.. Just SAD&pathetic!..  
(vLTMT500, VICTIM, 2020-02, 142)

Example 53 and Example 54 both cast aspersions on victims who have spoken out. In the former the writer presents themselves as supportive but views victims as mis-using the #MeToo movement to *demean and defame* others to build popularity. The second example demonstrates a reversal foregrounding the accused while counter-accusing Giuffre and her lawyers, similar to that described by Jones et al in relation to #NotAllMen (2022, p. 6). Repositioning the accused as now the victim of *false accusations* resulting from a "shakedown scheme" devised with her lawyers.

Example 53 I personally believe that this movement is essential but VICTIMS are not using it for the right purposes but in fact, accusing others of rape or sexual exploitation to demean and defame them and to be more popular in the media industry.  
Gulf News (vLNMT500, VICTIM, 2019-10, 76)

Example 54 He says he not only never had sex with Giuffre but that he has never met her. This makes him, he argues, a major victim of the #MeToo movement. I am a victim of her false accusations and I'm speaking out," he says. VICTIMS should speak out. Dershowitz believes that Giuffre concocted the allegations against him as part of a "shakedown scheme" devised with her lawyers.  
Sydney Morning Herald (vLNMT500, VICTIM, 2019-09, 186)

Both of these concordances invoke the resistant rape myth that women use false rape accusations to gain fame and fortune. News media may be over-reliant on clichés that reflect popular but erroneous beliefs due to commercial pressures by (Barca, 2018, p. 268).

The next section discusses the single concordance of a negative Activation Consequences theme.

#### **4.5.1.2 Consequences**

The only concordance expressing the consequences of speaking out comes from the vLNMT500 VICTIM dataset, and is given in Example 55. The phrase *take advantage* often carries negativity, suggesting "to seize an opportunity of unfairly profiting by a person or thing, esp. sexually." (Oxford English Dictionary [electronic resource], n.d.-L). Here, the writer describes the expectation that by self-declaring as a *victim*, the individual becomes a *hero*. However, the second sentence inverts the logic; because the *victim* hasn't forgiven, they can't be heroes, and therefore they aren't victims either. By this logic perpetrators would

only be called-out by victims who forgive them. By raising a legal case a victim has clearly not forgiven and therefore is not a hero, and therefore not a victim. Van Dijk suggests that victims acting against an expectation to forgive their perpetrators, run the risk of being scapegoated (2009, pp. 14, 25).

Example 55 And they take advantage now to say ‘Yes, I’ve been a VICTIM’, and they become heroes. A hero, for me, would forgive.

thetimes.co.uk (vLNMT500, VICTIM, 2020-04, 138)

#### 4.5.1.3 “Playing victim”

Within the full 24 month sexual violence corpora, the verb *play* only collocates with VICTIM and never with SURVIVOR, and never in NexisSH, refer to Table 43. According to a SketchEngine WordSketch the TwitterMT corpus has 44 instances of *play the victim* (LogDice 9.3), while TwitterSH has 20 (LogDice 8.4), and NexisMT has five (LogDice 6.9). As a comparison with general internet text, the EnTenTen21 returns *play* as a collocate of VICTIM with a LogDice of only 4.4, indicating that levels within this study are unusually high. All occurrences of *play the victim* are negative.

*Playing the victim* is a counter-accusation, and where there is confusion over who is the “real” victim we could ask “Which positions of power do these selves speak from? What kinds of benefits do their claims accrue to them and the communities they bring together?” (Chouliaraki, 2024, p. 36).

Returning to the coded vL500 datasets, Example 56 presents as a fairly measured case, with the initial *Umm* adding a tone of thoughtfulness and considered reasoning. However, the use of *plenty of women* suggests widespread false accusations of assault for revenge or financial gain, for which there is little proof. Furthermore, the trope *playing the victim*, is recognised as a rape myth (Appendix A – Rape Myths). The phrase *pretend to be* was also taken as analogous to *play*.

Example 56 Umm.. that there’s a possibility of everything actually being consensual and that the woman is playing the victim just to get money? Believe it or not, there are plenty of women who pretend to be VICTIMS of sexual harassment/assault to get money out of men. (vLTSH500, VICTIM, 2019-12, 52)

The opening *Spot on!* suggests the writer of Example 57 is agreeing with an earlier post in the same vein. The individual(s) referred to are accused of falsely claiming victim status. The phrase *playing the victim game* suggests that this is a commonly understood strategy.

Corpora	KWIC	Collocation with PLAY LogDice (freq)
TwitterSH	<i>victim</i>	8.4 (20)
	<i>survivor</i>	0
TwitterMT	<i>victim</i>	9.3 (44)
	<i>survivor</i>	0
NexisSH	<i>victim</i>	0
	<i>survivor</i>	0
NexisMT	<i>victim</i>	6.9 (5)
	<i>survivor</i>	0
EnTenTen21	<i>victim</i>	4.4 (N/A)
	<i>survivor</i>	0

Table 43, Collocation of PLAY with *victim* and *survivor* within the sexual violence corpora.

Example 57 Spot on ! They also abuse being a VICTIM (playing the victim game)Sexual-harassment claims , rape etc. (vLTSH500, VICTIM, 2019-07, 213)

The reference to a *fast track* in Example 58 also suggests that *playing the victim* is an established career pathway that by-passes the barriers that others have to work to overcome. The *we'll never know* phrase casts doubt on all celebrity MeToo victims. Debate on such quid pro quo transactions positions women lower down the career ladder as agentic but manipulative. An accused who is sufficiently senior and powerful to grant the *career success* is positioned as passive and forced into accepting sexual bribery.

Example 58 Playing the VICTIM is now a fast track to career success; We'll never know how many celebrity MeToo VICTIMS were fraudulentnt  
The Daily Telegraph (London) (vLNMT500, VICTIM, 2019-04, 255)

Failure to convict due to insufficient evidence is often referred to in these corpora as a deliberate attempt to defraud. Providing sufficient evidence is a constant difficulty for crimes that typically take place in private, and between people who are known to each other, and consequently continue to have to work together.

#### **4.5.2 Appraisalment**

Evaluation and judgement lies at the core of the Appraisalment social actor type. The evaluations of victims and survivors in these corpora are strongly polarised into “real” and bogus. Where an accuser is evaluated as bogus, there is often an accompanying counter-accusation of harm caused to another, unspecified, “real” victim or survivor. Table 41 illustrated that negative Appraisalment social Actor types are clustered within vLTSH500 VICTIM, and vLTMT500 VICTIM and SURVIVOR datasets. Table 44 shows the adjectives found within Appraisalments and their impact on the victimised, as identified in a manual exercise.

Sexual violence and the gender pay gap  
 4 Prosody and sexual violence

Dataset	KWIC	Total Negative Appraisements	Adjectives	Harm to “real” victims / survivors
vLTSH500	VICTIM	12	<i>real x 6</i> <i>genuine</i> <i>actual x 2</i> <i>true</i>	<i>taking from</i> <i>marginalises</i> <i>aren’t taken seriously</i> <i>crying</i> <i>diminish</i> <i>making it worse for</i> <i>hurt</i> <i>nullified</i>
	SURVIVOR	2	<i>true</i>	<i>won’t believe</i>
vLTMT500	VICTIM	55	<i>real x 18</i> <i>actual x 6</i> <i>true x 2</i> <i>professional</i> <i>fake</i>	<i>making it harder</i> <i>harm (3)</i> <i>hurt</i> <i>make ... into a circus</i> <i>causing damage</i> <i>disservice</i> <i>devalue</i> <i>detriment</i> <i>drowns voices</i> <i>delay justice</i> <i>questioned experience</i> <i>doesn’t help</i> <i>sacrificed</i> <i>undermines</i> <i>suffer</i> <i>made a mockery</i>
	SURVIVOR	10	<i>true x 2</i> <i>real x 2</i> <i>actual</i>	<i>devalue</i> <i>destroying the credibility</i> <i>hurt</i> <i>no longer believed</i>
vLN500	VICTIM	0	0	0
	SURVIVOR	0	0	0
vLNMT500	VICTIM	3	<i>alleged</i> <i>so-called</i> <i>genuine</i>	<i>taking a step back</i>
	SURVIVOR	0	0	0

Table 44, The most common evaluative adjectives for negative Appraisements, and the claimed harm to victims/survivors within the vL500 datasets.

Across the datasets *Real* (26) dominates the VICTIM Appraisal social actors followed by *actual* (8), *true* (3), and *genuine* (2), with all others having only single occurrences. The highest levels of evaluation for VICTIM are seen within vLTSH500 (75%), and vLTMT500 (32%), and for SURVIVOR within vLTMT500 (40%) and vLNMT500 (30%). Bogus victims/survivors are accused of a wide range of negative harms against anonymous but deserving victims/survivors.

The intended addressee for Example 59 is accused of turning #MeToo into a *circus*, and causing both *hurt* and *harm* to *real* but unspecified victims. Addressed to *you*, any reader might stop to consider whether they would be judged as *real* or bogus if they spoke out.

Additionally, the *you need help* and *extreme obsession* is a sarcastic questioning of the mental health of addressee.

Example 59 You make the #MeToo movement into a circus which only hurts the real victims You should be ashamed of the harm you're doing to REAL VICTIMS Get some psychological help ASAP you need help curbing your extreme obsession with this. (vLTMT500, VICTIM 2019-10, 78)

Other appraisals are shorter, and state their evaluation more directly. For instance, the writer of Example 60 first establishes themselves as compassionate and supportive with *victims need to be herd*, before casting judgement *these women are not victims*.

Example 60 I completely concur here... VICTIMS need to be herd...these women are not victims.. (vLTMT500, VICTIM 2019-04, 43)

The victim in Example 61, Christine Blasey-Ford, is judged for failing to conform to the Ideal Victim stereotype in the direct statement *not a innocent little VICTIM*<sup>67</sup>. While this could apply to any aspect of her, *if you get my meaning* gives it a seedy tone.

Example 61 #Christine was not a innocent little VICTIM , if you get my meaning. #MeToo #Kavanaugh #MondayThoughts Trump team held back dirt on Kavanaugh accuser, new book says <https://t.co/kxflwCdBVu> (vLTMT500, VICTIM 2019-07, 191)

While SURVIVOR examples seem generally kinder than the VICTIM concordances, as illustrated in Example 62, they still retain the pattern of a bogus victim harming *real survivors*. Many of the tweets in this study are driven by US politics, and here a supporter (Democrat) is being accused of making a false accusation in order to benefit Bernie Sanders, at the cost of *real survivors*. *She* is accused of *devaluing* and [mis-]using the #MeToo movement, before being judged as being *no victim*. In contrast the wronged *real survivors* are harmed by having their support movement *devalued* [for] *political reasons*.

Example 62 Enough with the self-righteousness. I say it's disgusting how she's using the #MeToo movement t advance the hit job started from the Bernie backers. She is devaluing the movement & real SURVIVORS for political reasons. She's no victim just a Bernie "supporter" (vLTMT500, SURVIVOR, 2019-04, 208)

This reflects Palomino Manjón's finding that #MeToo and accusations are tools intended to damage the GoP<sup>68</sup> (2024, p. 119).

In Example 63 the location of *survivors* between #*silencebreakers* and *WARRIORS* feels significant as it sets up an equivalence, foregrounding an ideological evaluation of the social actors (Bouvier, 2022, p. 186). This may set up an expectation that survivors will step forward, break the silence, and challenge with warrior-like courage, emphasised by the capitalisation. The *WARRIORS* are contrast with the diminished #*HollywoodHypocrites* who are *parading around*.

---

<sup>67</sup> Refer to Appendix A – Rape Myths.

<sup>68</sup> The US Grand old Party (GoP) or Republicans.

Example 63 Thank U John so much for this post U bet tomorrow night our hearts n' souls are very much w/ very REAL #Metoo #Kids2 #ChildVictimsAct #silencebreakers SURVIVORS ? WARRIORS while #HollywoodHypocrites parade around on the #AcademyAwards carpet ... ..and totally should be boycotted  
(vLTMT500, SURVIVOR, 2020-02, 434)

The single negative Appraisal concordance for vLTSH500 SURVIVOR, Example 64, is quite different in tone. Rather than a direct accusation, here there is a little room for uncertainty; *I hope these women are telling the truth, and if you did all these things*. However, the last phrase *we will find out* is somewhat menacing, suggesting unspoken ways and means.

Example 64 I hope these women are telling the truth because if not...ppl won't believe true SURVIVORS like me #metoo #MarilynManson if you did all these things you should pay only you know the truth and if you are guilty we will find out  
(vLTSH500, SURVIVOR, 2021-02, 113)

In an extremely judgmental tweet, the writer of Example 65 addresses a specific but anonymous third party, who is evaluated as a bogus victim who *nullifies the stories of real victims*. This results in the unspecified *real victims* being disbelieved by family and the police. It is implied that the person concerned won their court case because *Putting someone in jail*, suggests that there was evidence and a court believed them. The final sentence is a rhetorical question intended to make the accuser feel uncomfortable.

Example 65 Falsely accusing someone of sexual assault/harassment almost always nullifies the stories of real VICTIMS . The police don't believe us, neither do our families. Putting someone in jail, just to get back at them over something unrelated is beyond evil. How do you sleep at night??  
(vLTSH500, VICTIM, 2019-09, 166)

The three vLNMT500, VICTIM concordances are all reporting court cases, and all three journalists report prosecution accounts as unconvincing. In Example 66 the judge speculates that the accuser's *vivid imagination* was inspired by *allegations against Mr Weinstein and the MeToo movement*. This contrasts with the positive concordances (such as Example 25, Example 36, and Example 63) that praise the bravery of those speaking out and inspiring others to do the same. It is also strange because Weinstein had been found guilty a year earlier in February 2020.

Example 66 Mr Rimmer acquitted the multimillionaire at Westminster Magistrates' Court on Thursday and said the alleged VICTIM's account was 'riddled with troubling inconsistencies'. While commenting on the accuser's complaint, he said: "Might it be down to her vivid imagination Inspired by the allegations against Mr Weinstein and the MeToo movement. I cannot discount that it might.  
mirror.co.uk (vLNMT500, VICTIM, 2021-03, 82)

Within the manually coded vL500 datasets the Appraisal tweets were surprisingly negative. However, while writing Chapter 7, a WordSketch based on the full 24 month corpora (Table 70) confirming this negativity. The use of positive sounding adjectives, such as *real*, *actual* and *true* used to mark other unspecified victims, are also found in the pay gap datasets, and this type of representation is discussed in section 6.4.

The next section concerns itself with the Indifferentiation social actor type.

### 4.5.3 Indifferentiation

The negative concordances have much lower frequencies of Indifferentiation social actors than the positive datasets, and these collocate only with SURVIVOR. The Indifferentiation social actors coded as negative are generally expressing solidarity in order to increase their authority to speak for or about others. Often these tweets were written about high profile individuals while on-going court cases were being covered in the news media.

The phrase *as a sexual assault survivor myself* is used in Example 67 to claim authority to comment on those speaking out about the subject. A person advocating a zero tolerance<sup>69</sup> approach might consider that it is sexual harassment to squeeze or attempt to kiss a colleague. Consequently, having *her waist squeezed* would necessitate it being called out. However, the squeezed colleague might be reluctant to complain if they thought the local culture would be *enraged* by them doing so.

Example 67 As a sexual assault SURVIVOR myself, I felt undermined and enraged by the woman who had her waist squeezed crying harassment, and the woman who felt he ~might~ have kissed her hello had she not ducked, and the Hannity pal... (vLTSH500, SURVIVOR, 2019-09, 72)

The writer of Example 68 uses *we* invoking unity, but it isn't clear whether this is unity with *Vancouver*, or *'every other survivor'*. The message addresses Alyssa Milano and accuses her of failing to support *'every other survivor'*. While #MeToo was started as a safe space for victims to speak out, here it is positioned as coming with a responsibility to support and comment on all accusations. This suggests an obligation to speak out, discussed again in section 7.4.

Example 68 Hey @Alyssa\_Milano dont come back to Vancouver, Canada to work on #charmed! #CharmedCW we don't want your hypocritical self-righteous crap here anymore. YOU should have supported 'every other SURVIVOR' in the first place. You're a liar like the rest of them #MeToo (vLTMT500, SURVIVOR, 2020-04, 304)

In the context of Example 69 the use of *Please* suggests "incredulity or exasperation: 'for goodness sake', or 'come off it'." (Oxford English Dictionary [electronic resource], n.d.-x). The writer creates a sense of solidarity with *Many of us are survivors*, but then differentiates the *real survivors* from *Most of these women*. The use of *most* may lead readers to think that the bar for a qualifying claim is so high that *most women* do not reach it. Alongside that, the risk of *making it hard for real survivors* could deter an individual from speaking out themselves, not wanting to risk hurting others.

---

<sup>69</sup> Zero tolerance. "Complete lack of tolerance ... typically enforced by strict and uncompromising application of the law; a stated policy of this kind, designed to eliminate such behaviour." (Oxford English Dictionary [electronic resource], n.d.-P).

Example 69 Please. Many of us are SURVIVORS, yet that doesn't mean everyone who claims abuse has actually been abused. Most of these women are making it hard for real survivors . The #MeToo movement has turned into a joke.

(vLTMT500, SURVIVOR, 2020-05, 156)

Oddly, the solidarity of Indifferentiation is often juxtaposed with Differentiation used to emphasise separation from an Othered and discredited individual or group, discussed further in section 4.6.1. Indifferentiation appears to function in negative concordances to establish a writer's authority to judge and speak for others. The AS A VICTIM/SURVIVOR construct is explored further in 4.6.2. The next section discusses Passivation social actor types.

#### 4.5.4 Passivations

The Passivations Beneficialisation and Subjection are discussed together because of their close relationship, previously discussed in section 4.4.4.

Whereas the positive Passivations were more frequent for SURVIVOR, negative representations of both Subjection and Beneficialisation are more frequently found within VICTIM datasets. Around half of these occur in the vLTMT500 dataset. Table 45 shows the frequencies of Passivation concordances.

Corpora	KWIC	Beneficialisation	SURVIVOR representations as a % of VICTIM representations	Subjection	SURVIVOR representations as a % of VICTIM representations
vLTSH500	VICTIM	5		6	
	SURVIVOR	3	60%	5	83%
vLTMT500	VICTIM	10		21	
	SURVIVOR	5	50%	5	24%
vLN500	VICTIM	1		2	
	SURVIVOR	0	-	0	-
vLNMT500	VICTIM	6		9	
	SURVIVOR	0	-	1	11%

Table 45, Frequencies of negative Passivations for VICTIM and SURVIVOR within the positive vL500 datasets.

Manual coding of the datasets identified the phrases passivating VICTIM and SURVIVOR, and these are given in Table 46.

vLTSH500, SURVIVOR (8)	vLTSH500, VICTIM (11)	vLTMT500, SURVIVOR (10)	vLTMT500 VICTIM (31)
<i>an insult to survivors</i>	<i>victim-blaming Flores</i>	<i>a shitstain on survivors</i>	<i>believe victims</i>
<i>believe survivors</i>	<i>hurting victims</i>	<i>bar survivors from</i>	<i>black mail victims</i>
<i>harms real survivors</i>	<i>men fall victim</i>	<i>believe survivors</i>	<i>victims will be Blamed</i>
<i>hurting survivors</i>	<i>paid settlements</i>	<i>damage survivors credibility</i>	<i>don't care about/believe victims</i>
<i>mocking ... survivors</i>	<i>a lousy spokesperson for ... victims</i>	<i>not represent survivors</i>	<i>victims driven underground</i>
<i>survivors have been threatened</i>	<i>get money out of men</i>	<i>harmed survivors</i>	<i>exposing victims</i>
	<i>taking away from REAL victims</i>	<i>lied to, cheated, belittled survivors</i>	<i>victims of False allegations</i>
	<i>falsely calling them victims</i>	<i>stole from survivors</i>	<i>victims Get attention</i>
	<i>a victim... in shock</i>		<i>victims Let sex happen</i>
			<i>victims made to Suffer</i>
			<i>make a mockery of victims</i>
			<i>make it harder for victims</i>
			<i>victims of the #MeToo era</i>
			<i>plans laid out to victims</i>
			<i>shaming victims</i>
			<i>wronged</i>

Table 46, Manually identified phrases Passivating negative VICTIM and SURVIVOR within Twitter vL500 dataset.

Survivors, in Example 70, are both negatively Beneficialised by the failure of Kamala Harris<sup>70</sup> to represent them, and also Subjected as a subset of the American women receiving a *slap in the face*.

Example 70 Kamila Harris does not represent #metoo She does not represent sexual assault SURVIVORS , which are about ¼ women in America. She is a slap in the face to women across America. She's a fucking joke.

(vLTSH500, SURVIVOR, 2020-08, 19)

While the KWIC in Example 71 is SURVIVOR, *victims* seems to be used as a synonym. This was coded as Subjection because the *victims* are made a *laughing stock*, but the adjective *real* also made it Appraisement.

<sup>70</sup> Kamala Harris was Joe Biden's Vice President, and Democrat presidential candidate in 2024.

Example 71 There's nothing brave about making a laughingstock of real victims of sexual harassment and abuse. She's an insult to SURVIVORS.

(vLTSH500, SURVIVOR, 2019-04, 64)

In Example 72 Rose McGowan, an early #MeToo Weinstein accuser, is berated for failing to act, with the accusation that she *hasn't done a thing for* SURVIVORS. This again suggests an obligation to speak out, discussed again in section 7.4. *Riding the HW train for fame* refers to the trope that women accuse men falsely to gain benefit.

Example 72 Twitter hiatus..come back, Rose is still riding the HW train for fame,still hasn't done a thing for SURVIVORS and STILL taking credit for all things MeToo. And she still has delusional followers who don't care to look at facts. Amazing..

(vLTSH500, SURVIVOR, 2021-01, 463)

The VICTIMS in Example 73 are Beneficialised by the *hurt* incurred when they *lose their credibility*. This concordance is formatted as a question, but it is an oblique way of claiming the intended audience has made a false accusation.

Example 73 Did you know that if you falsely accuse someone of sexual assault, you're: A. Hurting an innocent person. B. Hurting the case. C. Hurting actual victims by making them lose their credibility. I can't talk about harassment and support VICTIMS w/o saying this. Do with it as u may.

(vLTSH500, VICTIM, 2020-07, 5)

The theme of victims falling short of the stereotype reappears in Example 74 where Alyssa Milano, like Rose McGowan in Example 72, is berated for failing to report Weinstein earlier. The concordance is coded as Subjection because the VICTIMS were *paid settlements* by Congress. Additionally, victims are Beneficialised by Congress protecting their names.

Example 74 Well...why didn't @Alyssa\_Milano report WEI STEIN when she knew all this crap was going on? Why hasn't Congress exposed names of sexual harassment VICTIMS of Congress who were paid settlements from secret slush fund funded by taxpayers...\$17MILLION?

(vLTSH500, VICTIM, 2019-07, 309)

Negative Passivations seem to be associated with unmet expectations to support and protect certain individuals, or groups, who are deserving but unspecified.

#### **4.6 Agency of SURVIVOR as a double-edged sword**

It is problematic that a component of the Ideal Victim (Sloan et al., 2018, pp. 11-22) is weakness, especially in an individualistic culture. Consequently, it may be that the victimised seek to be called something else in order to avoid the denigrated aspects of VICTIM, rather than seek SURVIVOR (Fohring, 2018, p. 162; Sehgel, 2016, p. 2 (in pdf reader)). Demonstrating agency, provides evidence that the victimised person is not weak and therefore not a victim. Hence, agency is a key issue for those claiming or given the SURVIVOR label.

In addressing agency, analysis focuses on the Activation social actor type, which is defined by the subject's action. Within the positive Activation social actor types there was a higher frequency of SURVIVOR (382) than VICTIM (290) concordances, while for negative Activation social actors the frequency is higher for VICTIM (25) than for SURVIVOR (6). This suggests that SURVIVOR is represented more frequently as positive and active than VICTIM. However, all

active verbs are not equal, with differences in the strength and also the nature of the implied action. For example, there is a difference between the actions of posting a short message on social media, and taking court action. The positive SURVIVOR concordances, in Table 36, contain greater variety of communication verbs than VICTIM. In contrast, the negative Activation concordances (Table 42) contain only three SURVIVOR verbs, *demand*, *claims*, and *taking credit* while VICTIM returns contain 12 in the form of Speaking out, five Playing the victim type, and a single verb phrase for Consequences. The situation is complicated for negative concordances by the use of sarcasm and irony, which are not always easily identified in the short text of tweets. However the negative SURVIVOR concordances appear to be sarcastic, illustrated in Example 47 and Example 48.

Agency can have either positive or negative prosody, and negative examples can position accusers as manipulating powerful men to gain wealth and fame, or settle old scores. However, victims may prefer the kind of agency that would enable them to rebuild their lives, change society, and sue for justice. Perhaps the ultimate loss of agency is for a victim to have their experience denied, for instance with a statement such as *you are not a victim*. This can have a few meanings, and these are explored in section 4.6.1

#### 4.6.1 If not a victim, then what?

A theme appearing across a number of negative VICTIM social actor types is the judgement of individuals as bogus victims/survivors. Evaluation as bogus is often accompanied with accusations of harming anonymous others who are judged more worthy of sympathy. SURVIVORS are never judged to be bogus.

To pursue this further, I carried out a SketchEngine Concordance search on the full 2.4 M word corpora, using the search phrases NOT A VICTIM and NOT A SURVIVOR. Across all four sexual violence corpora there were a total of 53 NOT A VICTIM concordances with a range of alternative nouns. In contrast, the NOT A SURVIVOR search returned either none at all, or only single instances, and these were of a *whether or not a survivor wants to...* structure. This section focuses on the higher frequency NOT A VICTIM concordances. Table 47 shows the alternative nouns used where these were stated for NOT A VICTIM.

Themes of concordances with the KWIK NOT A VICTIM					
Corpora	Survivor	Perpetrator/ Predator	Statistic	No Victimisation	Total
TwitterSH	1		0	11	12
NexisSH	0	0	0	1	1
TwitterMT	5	8	1	15	29
NexisMT	7	1	0	2	10
Totals	13	9	1	29	53

Table 47, SketchEngine concordances for the KWIK NOT A VICTIM

Fifty five percent (29) of concordances disagree that victimisation has occurred, and 90% of those were from the Twitter corpora. From the 15 within TwitterMT five show aggressive negative prosody like that seen in Example 75. The abrupt *they didn't* seems to slam shut that sentence, before the bald *You are not a victim*, which is amplified by the ten question marks. In contrast, within the TwitterSH corpus, a total of 11 deny victimisation but less

aggressively that those from TwitterMT, Example 76 is the most aggressive. The addressee is accused of *insult[ing]* the *vulnerable* and *genuine*, but the *you are not a victim* is tempered by the *bit of an insult*, and *it was awkward but*, which acknowledges the discomfort of the individual.

Example 75 You thought the #MeToo movement would embrace your post, they didn't . You are NOT A VICTIM. ??????????. (TwitterMT, NOT A VICTIM, 2020-01)

Example 76 I think this is a bit of an insult to the vulnerable who face genuine sexual harassment in the work place and don't have a voice, a microphone, or wealth. It was awkward but you're NOT A VICTIM. (TwitterSH, NOT A VICTIM, 2019-10)

The NexisMT concordance in Example 77 refers to a consensual affair between White House intern Monica Lewinsky and, then (1998) president, Bill Clinton. The issue here was not criminality or consent, but the abuse of a power differential which placed on President Clinton a duty of care for a very junior employee. The concordance credits Ms Lewinsky with agency, but the wrong type of agency for a victim, and at the wrong time, *that moment of reciprocal lust*. However, this writer overlooks the key issue was the misuse of power. *She may have been thrown under a bus* expresses sympathy for Ms Lewinsky's treatment by President Clinton, but still positions her as a bogus victim, agentically, doing a *disservice to the true victims*.

Example 77 She may have been a victim of Kenneth Starr and the American judicial process, she may have been thrown under a bus by Clinton himself in his repeated denials of "sexual relations", but at that moment of reciprocal lust in the Oval Office, she was NOT A VICTIM, and to describe her as such does a disservice to the true victims of sexual impropriety. Sunday Independent (NexisMT, NOT A VICTIM, 2020-04)

Twenty five percent (13) of the 53 concordances recognise the claim of victimisation, but reposition the victim as a *survivor*. The seven from NexisMT, are voiced in the first person and emphatically state that they are a SURVIVOR not a VICTIM, as in Example 78. These seven are all quotes of JK Rowling strongly defending her earlier comments on the transgender debate. *Certainly* emphasises that there is an important distinction between VICTIM and SURVIVOR.

Example 78 I'm a survivor, certainly NOT A VICTIM. The trans row blew up this week after Rowling highlighted an online article that discussed "people who menstruate". National Post, Canada (NexisMT, NOT A VICTIM, 2020-06)

Four of the five TwitterMT concordances are also in the first person, and like JK Rowling they highlight the importance of the difference between VICTIM and SURVIVOR for the individuals concerned. Example 79 is typical of these, using capitals to emphasis the *NOT*.

Example 79 While I appreciate being asked to tell my story please stop referring to me as a Victim of Child Abuse - I am NOT A VICTIM - I am a Survivor (TwitterMT, NOT A VICTIM, 2019-09)

Of the 53, nine (17%) use a third-party reversal representing the accuser as a *perpetrator*, or *predator* rather than a victim. Eight of these are from TwitterMT, with seven naming public

figures: Amber Heard<sup>71</sup> (2), Jacob Blake<sup>72</sup> (2), Jennifer Acuri<sup>73</sup> (1), Katie Hill<sup>74</sup> (1), and Cardi B<sup>75</sup> (1). Appearing exclusively in the TwitterMT dataset, a *predator* when applied to a human suggests a “ruthlessly exploitative or rapacious individual” (Oxford English Dictionary [electronic resource], n.d.-z).

In Example 80 Katie Hill is positioned as a *predator* for abusing her power relationship with a junior, positioned as Hill’s prey, *who feared retribution*. The rhetorical question *If you believe in #metoo* addresses #MeToo supporters, who would have to answer in the affirmative, leading to the conclusion that Hill’s behaviour was predatory. Hill was not charged with any crime.

Example 80 So @KatieHill4CA is NOT A VICTIM. She engaged in an inappropriate personal relationship with a subordinate who feared retribution if she severed the relationship (per leaked texts). If you believe in #metoo then @KatieHill4CA was a predator. (TwitterMT, NOT A VICTIM, 2019-10)

In contrast, a *perpetrator* is always a human carrying out a crime (Oxford English Dictionary [electronic resource], n.d.-v). All occurrences of *perpetrator* are within the NexisMT dataset and all refer to Amber Heard. Example 81 comments on her ongoing court case. *His suit* refers to Johnny Depp’s legal counter-case of defamation against Heard, which he later won. Heard claimed that Depp had organised a vitriolic social media campaign (Peplow, 2022) that, alongside death threats, mocked her testimony undermining her credibility in court in the real world. Depp had 9.8 million supporters on Twitter, while Heard had 3.6 million.

Example 81 Ms. Heard is NOT A VICTIM of domestic abuse; she is the perpetrator, his suit alleges. MailOnline (Nexis\_MT, NOT A VICTIM, 2020-06)

The individuals referenced in these concordances seem to have transgressed twice, firstly claiming to be a victim when they are perceived as bogus, and secondly being the “real” perpetrator.

There was only a single instance of *statistic* which seems to be a rare light-hearted comment, because the first sentence is a quote from the cult movie *Galaxy Quest* (Parisot, 1999). It goes on to make the serious point that every vote counts.

---

<sup>71</sup> Amber Heard and Johnny Depp are US actors who were involved in court cases following their divorce, claiming of domestic abuse, and defamation.

<sup>72</sup> Jacob Blake was a black American man shot in the back by police officers who received high profile support on social media, but attracted counter claims because he was shot while breaching a protection order on his girl friend’s home following earlier domestic abuse.

<sup>73</sup> Jennifer Acuri is a US entrepreneur who had an affair with Boris Johnson while he was mayor of London, resulting in an investigation for misconduct in public office. The investigation was dropped.

<sup>74</sup> Katie Hill was a US Congresswoman who had to resign over an inappropriate relationship with a campaign staff member. During the scandal nude photographs of Hill were published in the Daily Mail. This latter was claimed to be revenge by her ex-husband. Hill was not charged with any crime.

<sup>75</sup> Cardi B is a US rap artist who admitted to drugging and robbing men in her early career as a stripper, provoking a strong backlash on social media.

Example 82 Never give up and never surrender! I'm NOT A VICTIM, I am a statistic! I will vote and I will do whatever I have to, to save my country, my freedom and my women's rights! #MeToo (Twitter\_MT, NOT A VICTIM, 2020-09)

Having explored *not a victim*, next specific statements claiming a VICTIM/SURVIVOR position are described.

#### 4.6.2 “Speaking as a victim/survivor”

Within the concordances claiming a VICTIM/SURVIVOR identity, were a few using speech verbs alongside the phrase *as a victim*. To investigate this further I ran a SketchEngine search on the phrases AS A VICTIM and AS A SURVIVOR within the full sexual violence 24-month corpora, refer to Table 48. The table also shows the results of the manual breakdown into first and third person voices. Second person instances occurred no more than once in a corpus, so these were omitted.

The highest frequencies are for TwitterMT corpus first person AS A SURVIVOR (61, 98%) while the lowest AS A SURVIVOR within TwitterMT in the third person (2, 3%). The highest frequency AS A VICTIM construct is within TwitterSH first person with (32, 63%), and the lowest within NexisMT first person with (3, 15%).

Within the smaller numbers of Nexis concordances, there were comparatively higher percentages of AS A VICTIM examples, and in turn these had higher percentages of third person statements. Both of these factors could be due to legal language in the greater coverage of court cases in the news media.

The concordances with SPEAKING AS A VICTIM/SURVIVOR are generally coded as van Leeuwen Indifferentiation social actor types for their solidarity, and are also discussed under sections 4.4.3, 4.5.3, and 5.4.3.

Corpora	AS A SURVIVOR (1 <sup>st</sup> , 3 <sup>rd</sup> person)	AS A VICTIM (1 <sup>st</sup> , 3 <sup>rd</sup> person)	Total (1 <sup>st</sup> , 3 <sup>rd</sup> person)
TwitterSH	23 (1 <sup>st</sup> 21 (91%), 3 <sup>rd</sup> 2 (9%))	51 (1 <sup>st</sup> 32 (63%), 3 <sup>rd</sup> 15 (29%))	74 (1 <sup>st</sup> 53, 3 <sup>rd</sup> 17)
NexisSH	7 (1 <sup>st</sup> 4 (57%), 3 <sup>rd</sup> 3 (43%))	11 (1 <sup>st</sup> 3 (27%), 3 <sup>rd</sup> 7 (64%))	18 (1 <sup>st</sup> 7, 3 <sup>rd</sup> 10)
TwitterMT	61 (1 <sup>st</sup> 60 (98%), 3 <sup>rd</sup> 2 (2%))	43 (1 <sup>st</sup> 24 (56%), 3 <sup>rd</sup> 17 (40%))	104 (1 <sup>st</sup> 84, 3 <sup>rd</sup> 19)
NexisMT	18 (1 <sup>st</sup> 1 (6%), 3 <sup>rd</sup> 10 (59%))	20 (1 <sup>st</sup> 3 (15%), 3 <sup>rd</sup> 13 (65%))	38 (1 <sup>st</sup> 4, 3 <sup>rd</sup> 26)
Twitter Tot.	84	94	178
Nexis Total	25	31	56
Total	109	125	234

Table 48, Frequency of AS A SURVIVOR and AS A VICTIM in SketchEngine in the 2.4 M word sexual violence corpora

Note: the numbers of 1<sup>st</sup> and 3<sup>rd</sup> persons do not equal overall totals because this could not be reliably identified for some concordances.

TwitterMT AS A SURVIVOR concordances often appear to be part of an ongoing debate, as demonstrated in Example 83, and written by someone claiming personal experience. This

writer tells others who disagree not to reply, in an attempt to shut down the debate. The opening sentence *I support a woman's right to choose* expresses explicit solidarity.

Example 83 I support a woman's right to choose. Her body, her choice. If you disagree, please unfollow me. Not interested in arguing or having a 'moral' debate. AS A SURVIVOR of sexual assault, rape, stalking and harassment, you are not going to change my mind. (TwitterMT, AS A SURVIVOR, 2019-05)

The 43 AS A VICTIM concordances read as more reflective conversations about personal feelings on victimhood. Several of the writers discuss discomfort with the label, for example: *I no longer see myself as a victim; I don't see myself as a victim though; and don't want to be seen as a victim*. Example 84 shows a slightly different twist with the status of the victim reversed, not by themselves but by a third party describing how Marie's interview morphs into interrogation *as a suspect*. This illustrates the level of distrust women have in the reporting system, fearing the process will be "too dangerous or emotionally upsetting", which may contribute to under-reporting of sexual violence crimes (Advance Pro Bono, 2021, p. 22).

Example 84 The focus of the investigation became her credibility. And instead of interviewing Marie AS A VICTIM, they began interrogating her as a suspect. #Unbelievable on @netflix is painful to watch, yet accurate. We must do better. #metoo (TwitterMT, AS A VICTIM, 2019-09)

TwitterSH AS A SURVIVOR concordances are predominantly straightforward first person Tweets commenting on on-going court cases, as in Example 85.

Example 85 AS A SURVIVOR of military rape, which occurs at a rate of 1 in 3 military members, I thank you and the participants. Rape and sexual harassment should have no place in our government or society. (TwitterSH, AS A SURVIVOR, 2020-02)

Similar patterns are seen in the 51 TwitterSH AS A VICTIM concordances. Example 86 starts with a phrase often used by perpetrators "*She could ruin my reputation/livelihood!!!*", which sets self-inflicted reputational damage as equivalent to sexual violence. The exasperated writer continues with an emotional summary of an experience of being a victim, detailing the intense focus on the victim's personal life. The concluding *But sure. YOU'RE scared* is sarcastic.

Example 86 "She could ruin my reputation/livelihood!!!" M8, coming forward AS A VICTIM of sexual harassment ruins YOUR OWN reputation. Victim blamers constantly deny your trauma, pick you apart, and fill your life with hatred just for speaking out. But sure. YOU'RE scared. (TwitterSH, AS A VICTIM, 2019-09)

The NexisMT concordances frequently focus on the difference between VICTIM or SURVIVOR resulting in the lemmas appearing together. Four of these refer to the difficulty individuals have in identifying as a victim, using the phrases: *refused to see herself as; doesn't want to see herself as; it took me a long time to make the personal journey to look at myself; it was hard for me to even say what happened to me and identify myself*. Example 87 is from a natural language researcher studying MeToo news articles, and it encapsulates a perception, commonly expressed in these tweets, '*Victim*' evokes pity; '*survivor*' evokes admiration.

Example 87 “The person is described AS A VICTIM or a survivor of sexual assault,” she said. The word “survivor” evokes more positive emotions and coping mechanisms, she said. ‘Victim’ evokes pity; ‘survivor’ evokes admiration.”

Pittsburgh Post-Gazette (NexisMT, AS A VICTIM, 2019-09)

Another survivor, in Example 88, articulates a similar sentiment stating that *victim* is associated with weakness and the need for help. By starting two sentences with the emphatic *No, I did not need to*, they emphasise their agency, but find themselves forced *into victimhood*. The rhetorical question in the final sentence makes clear the reality of having to accept victimhood in order to take *disciplinary action*. The writer expresses that the reality of the Title IX legislation is that to be a *survivor*, they have to concede something important.

Example 88 But the new Title IX policies do not acknowledge my identity AS A SURVIVOR . They force me into victimhood. Survivors should not be punished for their strength. No, I did not need to go on academic leave. No, I did not need to go to the hospital. Does it qualify me less for pursuing disciplinary action against my assaulter?

The Georgetown Voice: Georgetown University  
(NexisMT, AS A SURVIVOR , 2020-05)

The lowest frequency of the search phrases were found in NexisSH, and these few tended to focus on the consequences of using the VICTIM label. This in theory should provide support for the individual, but in practice may have direct negative implications on a career, as expressed in Example 89. Although the writer is the wronged party, *someone who causes problems* suggests that other employers might see the problem as the victim doing the reporting rather than the victimisation. *What that would mean of getting hired again* implies that the writer fears not be able to find work again.

Example 89 I don’t want to be seen AS A VICTIM and I also would have fear of what that would mean of getting hired again because I don’t want to be seen as someone who causes problems. The Australian (NexisSH, AS A VICTIM,2021-02)

Dame Jasvinder Sanghera, in Example 90, is recounting her decision whether to pursue a sexual harassment claim against a member of the UK House of Lords. She alludes to the difficulties of living with the negative semantic prosody attached to the VICTIM label previously, and now *faced with having to accept it again*, making clear her reluctance to do so.

Example 90 I also had my personal struggles to deal with. AS A SURVIVOR of a forced marriage, it had taken me years to rid myself of the “victim” label and now I was faced with having to accept it again.

The Independent (NexisSH, AS A SURVIVOR, 2019-10)

In contrast to the accusations of VICTIMS *riding the HW train for fame*, referred to in Example 72, these concordances suggest that the consequences of speaking out as a victim are generally negative and not taken on lightly.

## 4.7 Conclusion

This chapter is aligned to research questions 1A and 1B, and it has described the analysis of negative and positive representations of VICTIM and SURVIVOR to identify which was more frequently negative.

The answer to RQ1A is yes, VICTIM representations are more frequently coded negative than SURVIVOR representations.

The majority of all concordances in the vL500 datasets are coded positive. However, within the vL500 datasets VICTIM is coded as negative more frequently than SURVIVOR. The range of negative concordances varies between 1% for vLNSH500, and 18% for vLTMT500. The SURVIVOR datasets have lower rates of negative coding, between 0% for vLNSH500 and vLNMT500, and 5% for vLTMT500. Hence, the negative representations that are present are skewed towards the VICTIM concordances. The negativity within vLTMT500 illustrates how challenge can be enacted through a #hashtag, opposing its original purpose (Jones et al., 2022, p. 9). This study has heavily deduplicated tweets in order to include as many different voices as possible, meaning that these findings cannot be related to the experience that a user would have had viewing Twitter at the time this data was harvested. Because of the bias of Twitter's algorithms toward negative and extreme tweets, discussed in section 3.2.1.2, users at the time would have seen much larger numbers of negative, and similar, postings.

In answer to RQ1B, the main themes for VICTIM are those mapping to the social actor types: Activation, Appraisalment, Beneficialisation, Indifferentiation, and Subjection; and also the sub-themes of Activation: Blaming/shaming, Consequences, Playing victim, Speaking out, and Victim vs survivor. SURVIVOR shares all of those themes, and with greater numbers than VICTIM in: Activation, Beneficialisation, and Indifferentiation. The Passivations Beneficialisation and Subjection are the most common social actor types, with Subjection always the more frequent for VICTIM, and Beneficialisation more frequent for SURVIVOR. Both Passivations can feature individuals as receiving or subject to positive expressions of support, however victim concordances are more aggressively adversarial in their advocacy. Appraisalment concordances, dominated by VICTIM, are usually negative evaluations, appearing in positive concordances as calling out unacceptable behaviour. Appraisalments usually include strongly polarised statements suggesting little uncertainty. Although present only in low frequencies, both VICTIM and SURVIVOR Indifferentiation examples are notable in their use of claiming authority to speak for all SURVIVORS, and often to criticise or diminish another person speaking out. Even within the same theme SURVIVOR concordances have a kinder and more personal tone than VICTIM examples. During the analysis it appeared that VICTIM is denigrated by various means that are not applied to SURVIVOR. For example, those subject to sexual violence may be *victim-blamed*, or accused of *playing victim*, but never *survivor-blamed* or *playing survivor*. Additionally, no SURVIVOR equivalent was found to mirror the phrase NOT A VICTIM. Taken together, the evidence of this chapter suggests that VICTIM is more strongly associated with negative prosody than SURVIVOR.

The answer to RQ1C is that when they claim the SURVIVOR label, based on the analysis in this chapter, the disadvantaged appear to gain a more beneficial identity. That SURVIVOR is associated with less negativity suggests that it is culturally preferred over VICTIM. Based on this first analysis chapter the evidence suggests that individuals appear to benefit from

adopting the SURVIVOR label. SURVIVOR is associated with less negativity, and is more frequently associated with Activations such as Speaking out, and the calling out of unacceptable behaviour. Additionally, SURVIVOR is never exposed to accusations such as “playing the survivor” or “playing the survivor card”. The analyses in the following chapters, especially Chapter 7, revisit this provisional answer to RQ1C, as they situate it within the wider context of potential personal costs and the implications for society.

## 5 Prosody and the pay gap

### 5.1 Introduction

The purpose of the analysis in this chapter is to confirm or discount the claims of research question 2, which is repeated below for convenience.

RQ2 How are those disadvantaged by the gender pay gap represented within tweets in relation to positive and negative affective responses, and how does this compare to representations associated with sexual harassment? This question is addressed in three parts:

RQ2A Do the pay gap datasets contain more negative representations than the sexual violence datasets? (section 5.3)

RQ2B What are the dominant themes associated with those disadvantaged by the pay gap? (section 5.4 & 5.5)

RQ2C How are the themes found in the pay gap datasets different from those of the sexual violence datasets? (sections 5.5 & 5.7)

Although the vLNPG500 dataset is not the main focus of RQ2A, it is analysed alongside vLTPG500 to help differentiate the influences of topic and platform, and to provide a comparison.

We speak of individuals being victimised by sexual harassment or assault, but we do not speak of anyone being pay gap-ed. In fact it is the phenomenon of the gender pay gap itself that is usually discussed with only indirect references, if any, to those impacted. This should not be surprising as being a statistical phenomenon, discussed in section 2.2.2, refers to populations rather than individuals. Identification of those disadvantaged by the pay gap, is consequently approached via the SketchEngine `PAY GAP IS A ...` construct. Where individuals or groups are directly discussed, it is most commonly in relation to equal pay, a component of the pay gap that affects individuals, but does not account for societal barriers.

In section 5.2 the coded pay gap datasets are analysed to identify positive and negative representations, leading into a discussion of the most prevalent van Leeuwen social actor representations in section 5.3. Then social actor types coded as positive are discussed in more depth in section 5.4, followed by those coded as negative in section 5.5.

What this chapter's analysis means for RQ2 is explained in section 5.7.

### 5.2 +/- prosody representations

The +/- prosody<sup>76</sup> flags were assigned to concordances according to whether a person impacted by the topic would interpret the prosody of a representation positively and feel comfortable speaking out, or negatively and be less likely to speak out.

---

<sup>76</sup> A full discussion of the definition of +/- prosody is given in section 3.4.2.

Sexual violence and the gender pay gap  
5 Prosody and the pay gap

Datasets	Positive	Negative	Neutral
vLTPG500	67%	21%	12%
vLNPG500	100%	0%	0%

Table 49, The +/- prosody composition in the vLPG500 datasets

The vLTPG500 dataset has both the highest level of negative concordances (21%) and also more neutral concordances (12%). In contrast, the vLNPG500 dataset has so few negative and neutral concordances that these rounded to zero (0%), illustrated in Table 49.

The mapping of +/- prosody onto the van Leeuwen social actor types<sup>77</sup> is explored in the next section, 5.3.

### 5.3 Van Leeuwen social actor representations

Coding of the van Leeuwen social actor types is described in section 3.4.1.

The +/- polarity codes are shown mapped to the van Leeuwen social actor types in Table 50.

Datasets	KWIC	Activation	Appraisement	Beneficialisation	Indifferentiation	Subjection	Suppression
+ vLTPG500	Pay Gap	16	4	174	9	10	91
Neutral vLTPG500	Pay Gap	4	0	12	0	1	27
- vLTPG500	Pay Gap	8	23	19	0	4	42
+ vLNPG500	Pay Gap	26	5	371	3	27	22
Neutral vLNPG500	Pay Gap	0	0	2	0	0	0
- vLNPG500	Pay Gap	0	0	0	0	0	0
Total vLTPG500	Pay Gap	28	27	205	9	15	160
Total vLNPG500	Pay Gap	26	5	373	3	27	22

Table 50, Highest frequency social actor representations  
Refer to footnote 76, page 135.

Beneficialisation concordances are predominantly positively coded, and account for the highest frequencies. The second most frequent social actor here is Suppression, which was absent from the sexual violence datasets. Both pay gap datasets also contain lower rates of positive codings for Activation, Appraisement, Indifferentiation, and Subjection social actor types, all at rates lower than 5%. The vLTPG500 dataset has negative codings for all of those social actor types except Indifferentiation, the social actor type associated with solidarity. There are no vLNPG500 concordances with negative prosody.

<sup>77</sup> van Leeuwen social actor types were introduced in section 3.4.

The selection of the social actor types for analysis (Activation, Appraisal, Beneficialisation, Indifferentiation, Subjection and Suppression) was described in section 4.3.3.

The next section, 5.4, discusses some vLPG500 examples coded with positive prosody, before turning to explore some negative examples in section 5.5. The neutral coded concordances are not analysed, because the focus of the study was on the higher frequencies of positive and negative concordances. A deeper investigation of neutral coded concordances may make interesting future research.

#### 5.4 Positive representations within pay gap datasets

The breakdown of positive representations in the pay gap datasets is presented in Table 51<sup>78</sup>. Note that 19 Nexis representations and 20 Twitter social actors were coded to social actors below the threshold for inclusion. Additionally, 166 Twitter representations were coded as negative and neutral.

Dataset (Total positive codes)	Activation (% of positives)	Appraisal (% of positives)	Beneficialisation (% of positives)	Indifferentiation (% of positives)	Subjection (% of positives)	Suppression (% of positives)
vLTPG500 (334)	16 (5%)	4 (1%)	174 (52%)	9 (3%)	10 (3%)	91 (27%)
vLNPG500 (500)	26 (5%)	5 (1%)	371 (80%)	2 (0%)	27 (5%)	22 (4%)

Table 51, Frequency of positive coded social actor types  
Refer to footnote 76, page 135.

Sections 5.4.1 – 5.4.5 now explore the van Leeuwen social actor types, drawing on examples from the vLNPG500 and vLTPG500 datasets. Unlike the sexual violence examples in Chapter 4, the people being discussed in these texts are often not referred to directly, so representations are then inferred.

##### 5.4.1 Activation

For the Activation social actors in both vLNPG500 and vLTPG500 datasets the same three themes predominate (Table 52). However, the Speaking Out theme for Nexis has four times more (14) than for Twitter (3).

<sup>78</sup> The positive figures in Table 51 plus the negative figures in Table 58 do not total 500 because of the concordances coded as neutral, and the social actor types that did not meet the threshold for analysis (10 concordances).

Dataset (total)	Themes within Activation				
	KWIC	Speaking out	Causes	Consequences	Others*
vLTPG500 (16)	Pay Gap	3	7	4	2
vLNPG500 (26)	Pay Gap	14	7	5	

Table 52, Most common themes within positive Activation social actor types within vLNPG500 and vLTPG500 datasets

\* "Others" includes the other themes with only single instances in these datasets.

Refer to footnote 76, page 135

Of the 16 vLTPG500 concordances 15 of the subjects are *women*, one is *moms*, and all are Genericised. The verbs and verb phrases attributed to these 16 include: *redesigning*; *talking*; *providing responses*; *congregating*; *constructing shanks*; *communicating*; *asking*; *negotiating*; *contributing*; [doing] *women's work*; *taking maternity leave*; *giving up so much*; *screaming*; *working harder*; and *making it up*.

Within the 26 vLNPG500 concordances the disadvantaged Genericised groups are more specifically identified, including: *civilian employees* (3); [BBC] *presenters* (2); *football/soccer/baseball players* (5); *athletes*; and *teachers*. The verbs and verb phrases attributed to them include: *stage information pickets* (3); *demanding action* (2); *calling out Hollywood*; [expressing] *fury*; *authored the report*; *protesting*; *striking*; *testifying*; *reaffirm stance*; *demonstrating*, and *dealing with the consequences*.

Note that many of the positive concordances are quoting and call out negative phrasing, in order to lend support. Activation's three main themes are now explored in the following sections.

#### 5.4.1.1 *Speaking out*

Speaking out is a broad theme characterised by the use of communication verbs or verb phrases. As news media articles, all of the vLNPG500 concordances theoretically could have been coded as Speaking out, but in-line with the research questions, focus is retained on those impacted by the pay gap who are represented as speaking out.

The writer of Example 91, in discussing an equal pay issue, is surprised that the cause of the pay gap is attributed to a failure to speak out effectively, phrased *women asking for less*, which would constitute victim-blaming, and suggests Erlich's (2001, p. 148) "culturally-sanctioned ignorance" that avoids recognition of an issue that society doesn't want to address.

Example 91 It surprised me just that it seems to indicate that the gap is purely based on women asking for less, but yes, the finding that there's a gender PAY GAP isn't surprising. (vLTPG500, 2019-04, 244)

The *purely* isolates the *asking for less* from the social context, such as awareness of colleagues salaries, which would provide a baseline. Reward Expectations Theory would explain that women may ask for lower salaries that are in line with those of other women they know, and when they don't know what their male colleagues earn (Auspurg et al., 2017). In a related issue in Example 92, a previous comment discounts an earlier tweet (not part of this dataset) locating the cause as a personal failing of women. Rather, it alludes to

constructed social barriers such as the enculturated stereotype of women as less assertive and direct.

Example 92 Smarter negotiating isn't enough, bc the pay gap isn't women's fault...it can't just be on women to fix this problem. A lot of us were raised that it's not polite to talk about \$. To close the PAY GAP , it's time to stop being polite and start talking about equal pay.? Issa Rae<sup>79</sup> (vLTPG500, 2019-11, 107)

Being *polite* is positioned as a disadvantage, and this points to another issue. A meta study showed that women negotiating were: judged as non-normative and less likable; held to harsher standards; and subject to greater push-back to accept lower offers than men (Cheng et al., 2018; PayScale, 2020). Cheng suggests that women's negotiation skills may not be effective until perceptions of women change. This situation is discussed further in sections 6.5.6.

The organisers of the *mass scream* apparently agree that this is not the time to be polite, providing this dramatic verb phrase for the headline promoted in Example 93.

To investigate the general use of SCREAM I viewed its collocates in the enTenTen21, and the top ten are given in Table 53.

Word	LogDice
<i>kicking</i>	8.09
<i>yelling</i>	7.59
<i>loud</i>	7.38
<i>crying</i>	7.31
<i>agony</i>	7.22
<i>Primal</i>	7.21
<i>heard</i>	6.85
<i>yell</i>	6.81
<i>shouting</i>	6.64
<i>bloody</i>	6.64

Table 53, Top ten collocates of SCREAM within the enTenTen21

The collocates are all strongly emotive, with the strongest collocate being *kicking* (8.09 LogDice). Others include *agony* (7.22 LogDice) and *bloody* (LogDice 6.64). *Bloody* most frequently appears as part of the phrase *screaming bloody murder* in the enTenTen21 (0.03 per million tokens). An unexpectedly strong collocate was *Primal* (7.21 LogDice) also listed as *primal* (LogDice 6.12) which, according to the Oxford English Dictionary, expresses "needs and fears held to form the origins of emotional life" (n.d.-A). With this deeper appreciation of *scream*, focus returns to Example 93.

The use of the agentic verb *stage* positions the act as a conscious and organised demonstration, and the prefix *mass* strengthens *scream* further. This event ties together the twin issues of *domestic violence* and the *pay gap*, suggesting a relationship.

Example 93 Women stage 'mass scream' in Switzerland over domestic violence and gender pay gap - The Guardian <https://t.co/xpAsclONJ6> (vLTPG500, 2020-10, 323)

While vLTPG500 examples often express personal opinions, vLNPG500 articles are typified by reporting of high profile figures such as Jennifer Lawrence's<sup>80</sup> response to discovering her own un-equal pay situation (Feeney, 2015). In Example 94 Lawrence is represented as writing

<sup>79</sup> Issa Rae is a US actor, writer and producer. She is also an activist for civil rights and women's rights.

<sup>80</sup> Jennifer Lawrence is a US actor who identifies as a feminist, and advocates for civil and women's rights.

*an impassioned essay* for a feminist newsletter following an e-mail hack. The reference to an *e-mail hack* suggests that ordinarily *Hollywood* salary information is kept secret, itself an issue when transparency is associated with lower pay gaps.

Example 94 After the Sony email hack revealed that she was paid less than her male American Hustle co-stars, Lawrence addressed Hollywood's gender PAY GAP issue in an impassioned essay for Lenny Letter, Lena Dunham and Jenni Konner's now-defunct newsletter. Evening Standard (vLNPG500, 2019-08, 187)

#### **5.4.1.2 Causes**

Concordances with a Causes theme focus on reasons for the pay gap's existence, commonly involving: women's care responsibilities; the consequent need for flexible working; unequal pay for similar work; and career segregation. Strangely, the same reasons are often also given by others as consequences of the pay gap. Assignment as Cause or Consequence themes is subjective, but the tone used in this polarised debate generally left little ambiguity.

Example 95 suggests that the pay gap is related to the burden of *unpaid care work*, and contrasts it with the *8 trillion* it is worth to the global GDP. The *oh yeh* introduces the final sentence as a casual afterthought, but the figure 37% is substantial, suggesting that this is meant as sarcasm.

Example 95 As per @CCriadoPerez excellent book Invisible Women, women's unpaid care work contributes 8 trillion to the annual global GDP, but they have less access to household finances. Oh yeh & the average gender pay gap is 37.% ... (vLTPG500, 2020-12, 263)

Unpaid care is often gendered by the possessive *women's* as in Example 96. Here its value is positioned as fundamental to the functioning of communities rather than financial. This writer states that it is *women's work* that keeps communities functioning., and juxtaposes that with the following sentence *Close the gender pay gap*. Unpaid care work and the pay gap are often linked implicitly, if not explicitly, with the UK government's 2019 "Gender Equality Roadmap" (Government Equalities Office, 2019, pp. 6-8) pointing to caring responsibilities (section 2.2.2). Fineman (1995) had called for a complete redefinition of "family" in recognising the difficulties women face in unpaid care work.

Example 96 was also coded as Speaking out for the emphatic *We need equal pay now!* which clearly states that the legislated 50 year old base-line of equal pay has yet to be achieved here.

Example 96 Without women's work our communities will NOT function. Close the gender PAY GAP @fordnation. We need equal pay now! (vLTPG500, 2020-04, 266)

The book Invisible Women (Criado-Perez, 2019), mentioned in Example 97, appears in 14 concordances of the TwitterPG corpus. In it Criado-Perez argues that *biology*, meaning motherhood, is used to obscure other *economic/social factors* that contribute to the pay gap. Women are represented as agentically *taking maternity leave*, by making a "choice-that-isn't-a-choice" (Criado-Perez, 2019, p. 76). By accepting the only option, they are represented as responsible for their reduced earnings.

Example 97 I have read *Invisible Women* and I agree on a fair bit of what you say. Where I disagree is that emphasising biology over economic/social factors leads to folk believing the gender PAY GAP is based on women taking maternity leave etc & not looking at issues like pay transparency. (vLTPG500, 2020-01, 277)

Women are represented as acting agentially in *concentrating in more junior roles* and more often using, *flexible work arrangements* in Example 98. However, these actions may result from the previously mentioned unpaid care work, rather than unrestricted choices. Women's implied failure to gain promotion out of the *junior roles* suggests a 'glass ceiling', conjuring an image of a physical, yet essentially invisible, barrier to career progression (Oxford English Dictionary [electronic resource], n.d.-p).

Example 98 While there is little difference in the base salaries of men and women in the public service, the gender PAY GAP is exacerbated by the concentration of women in more junior roles and their greater use of flexible work arrangements. Canberra Times (vLNPG500, 2019-11, 236)

The term *unadjusted*, discussed further in section 6.3, is used in US official reporting, and suggests a raw figure that needs to be adjusted. Other reporting, such as *Women's Rights in Review 30 Years After Beijing* (UN-Women (United Nations Entity for Gender Equality and the Empowerment of Women), 2025) and the *Global Gender Gap Report 2025* (World Economic Forum, 2025) does not use the adjusted/unadjusted terminology. Example 99 is from one of 74 articles from Singapore, and it explains that the adjustment discussed here is career segregation. While the statement men and women tend to work in different occupations is true, that does not address the issue that women face restricted choices in the workplace due to their need to accommodate unpaid work caring for children and others which requires flexibility (Bennett, 2024, p. 347; Shaw et al., 2018, pp. 3-5). *Tend to* suggests a constant bias towards an option (Oxford English Dictionary [electronic resource], n.d.-M), suggesting that the individual has a viable alternative.

Example 99 This unadjusted figure is often used in international comparisons by the likes of the Organisation for Economic Co-operation and Development (OECD). The 16.3 per cent PAY GAP is largely driven by the tendency of men and women to work in different occupations, the study said, as women tend to be in lower-paying jobs compared to men, who continue to be over-represented in higher-paying occupations.

MCN International (Singapore) (vLNPG500, 2020-01, 302)

### 5.4.1.3 Consequences

Concordances coded with the Consequences theme provide information or opinion on effects resulting from the pay gap. Some features are viewed as both cause and consequence by different writers, according to their viewpoint. Crucially, by flipping Cause to Consequence the issue becomes the responsibility of the impacted individual. Common topics discussed in this section include: women's sport (soccer, tennis, golf); being worth less; and the unpaid care work discussed in the previous section.

When the USA women's soccer team won the world cup in 2019 their pay was hotly debated in the media, as illustrated in Example 100. Figure 16 shows the tweet that was linked to Example 100, which has since been deleted.

This writer acknowledges that their funding was a consequence of the degree of interest in the women's game, but nevertheless suggests that their pay was unfairly low, capitalising *Problem* emphasises the point. However, the level of interest in the women's game itself may be a product of gendered media bias (Alkhamash, 2021, p. 24).

The writer evokes emotion with *giving up so much* and the patriotic *play for their countries* further highlights a sacrifice.

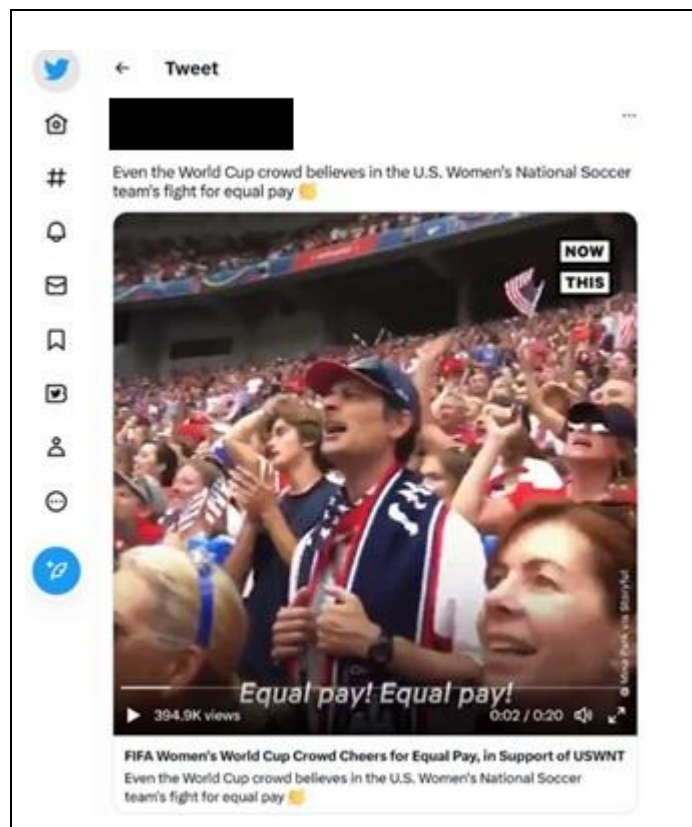
The tone of Figure 16 reflects Alkhamash's (2021, pp. 23, 27)

finding that the top three emotions in her positive tweets expressed love, pride and happiness in the team's achievements. Within her negative tweets she found strong elements of discontent with the US Soccer Federation (USSF) for institutionalised gender discrimination, and this disappointment is evident in Example 100.

Example 100 The PAY GAP needs to be addressed. I understand the men's games draw more money, but these women are giving up so much to play for their countries. It's a Problem throughout womens sports. (vLTPG500, 2019-04, 309)

Within a household unpaid care work and low pay may become self-reinforcing, with the lower paid partner giving up working hours to take on any required unpaid care work. The issue becomes an identifying characteristic when the lower paid/unpaid care worker partners are more often of a specific gender. In Example 101, support is expressed for women who have to *work harder just because of gender*, while they *contribute higher to the*

Figure 16, Tweet following USA women's soccer win  
Linked to Example 100 (now deleted).



*GDP*. This could be another reference to the figures quoted in *Invisible Women* (Criado-Perez, 2019) published the previous year. The final sentence starts *idk*<sup>81</sup> *man* suggesting perplexity at this ongoing problem, and finishes with *already*, originally an American Jewish intensifier expressing impatience or exasperation (Oxford English Dictionary [electronic resource], n.d.-b) with the two question marks adding further weight.

Example 101 this isn't really something to smile about. The fact that women have to work harder just because of gender and they still contribute higher to the GDP? idk man fix the gender PAY GAP already ?? (vLTPG500, 2020-08, 327)

Example 102 argues that austerity in the aftermath of the 2007 financial crash has disproportionately harmed individuals who occupy more than one disadvantaged demographic. A specific group of (*white*) *women*, made more pointed by the brackets, are positioned as selfish for being concerned only with *childcare expense*. They are called out because they *never makes a noise* to advocate for other groups. This privileged white subset is contrast with the greater disadvantage experienced by those already facing other barriers for instance because of: gender, race, disability, or single motherhood. The themes of family and unpaid care responsibilities have also been identified in German (2013) and Italian (2010) language pay gap focused corpora (Nadone, 2018, pp. 181, 182) suggesting the pervasive nature of this issue.

Example 102 The type of woman who only really notices sexism when it's parsed in terms of childcare expense and (*white*) women's PAY GAP but never makes a noise about how austerity has hit (for example) single disabled mums and women of colour (and those who are both those & more) hardest. (vLTPG500, 2019-11, 364)

The vLNPG500 Consequence concordances illustrate frustration at the gender pay gap and the damage it does to long term financial health. Example 103 combines two themes: the Consequence that women are *being peddled a myth*, and also Speaking out for the *sent a follow-up email*. The *myth* being peddled is the traditional stereotype that women are *too modest* and *too timid to raise issues*. As discussed for Example 92, it can be argued that lack of negotiating skills is not the problem.

Example 103 When I sent a follow-up email to ask what was being done about the company's considerable gender PAY GAP , it went ignored. We've long been peddled the myth that women are too timid to raise issues like this, too modest to make the case for a pay rise. This is b\*\*\*\*\*s.  
The Herald (Glasgow) (vLNPG500, 2020-10, 251)

In Example 104 the Daily Mirror grabs attention with the opening sentence *WOMEN need to work until 103*, which readers may guess is twenty years beyond the 83.1 years average female life expectancy (Office for National Statistics (ONS), 2020). The use of *persistent*, and *still* indicate impatience with this problem which will perpetuate the disadvantage into the future.

---

<sup>81</sup> Idk = I don't know.

Example 104 WOMEN need to work until 103 to match men’s pension pots, a report says. The proportion of women saving adequately for retirement has hit a new high of 59%. Yet the persistent gender PAY GAP and part-time working means women still save £1,300 a year less than men into pensions.

Daily Mirror (vLNPG500, 2020-11, 381)

**5.4.2 Appraisalment**

The positive examples of Appraisalment in the pay gap datasets are shown here as Table 54, and these are much less frequent than those within the sexual harassment datasets given as Table 44.

Themes within Appraisalment						
Dataset (total positive Appraisals)	KWIC	Speaking out	Causes	Consequences	Denial	Abuse
vLNPG500 (5)	PAY GAP	0	3	4	0	0

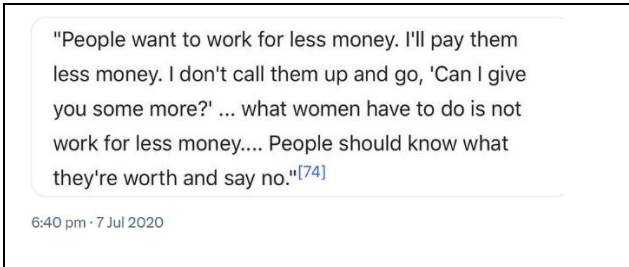
Table 54, Most common themes for positive Appraisalment social actor types  
 Refer to footnote 76, page 135.

Within vLTPG500, Appraisalment only ever exists for a concordance alongside another code. In both Example 105 and Example 106 the second code is Subjection. In Example 105 the writer uses *stop* to indicate that it is a current issue that women are made to *feel shit about themselves for the gender pay gap*.

Example 105 Understandable. Big Q is what do you do - stop making women feel shit about themselves for PAY GAP ? Force men to take more active role /penalize them? Virtue signal (what’s happening today) about patriarchy? Optimize over what time frame - years, decades, generations? (vLTPG500, 2020-11, 26)

Figure 17, Tweet “People want to work for less money”

Like Example 105, Example 106 is calling out an earlier tweet, shown as Figure 17, because it *blames women*. Being awarded higher pay simply by asking assumes that employers and employees have equal power and access to pay information.



Example 106 She also blames women in regards to the gender PAY GAP in Hollywood. (vLTPG500, 2020-07, 32)

The BBC equal pay scandal is discussed in Example 107, where the writer acknowledges the disparity, but also evaluates some BBC employees as *paid too much*. It's not clear whether *these people* refers to all BBC presenters, just the female ones, or possibly just Zoe Ball. Gary Linkeker's salary of £1.75 M released at the same time, is not raised.

Example 107 I believe Zoe Ball is the highest paid BBC female presenter on £370K. I have no doubt that there is a sex-dependent PAY GAP and also no doubt these people are paid too much. (vLTPG500, 2020-04, 135)

Women are also *blamed* for the pay gap within Nexis, with *firms* in Example 108 pointing to women's failure to progress their careers to *high paying jobs*. The phrases *then it emerged* represent as a revelation the highest claimed pay gap figure in this study, 70%.

Example 108 "Initially, when they started the compulsory reporting firms blamed their PAY GAP on the lack of women in highpaid jobs," Denis-Smith says "The gap was already huge, but then it emerged that many firms chose not to disclose the gap at the top." At partnership level, it later emerged, the gap in some cases was as high as 70 per cent. The Times (London) (vLNPG500, 06-2020, 461)

In Example 109, from Malta Today, Mamo alludes to structural barriers beyond equal pay that contribute to the gender pay gap. However, he suggests there is a *tendency of women to opt for lower-paying jobs*, with "opting" implying that another option is available when it may not be. *Child-rearing* Mamo refers to the family responsibilities for which Nardone (2018, p. 181) found evidence in both German and Italian corpora. Mamo does not suggest a reason why men are not opting for career breaks for child-rearing and lower paid roles.

Example 109 Mamo emphasised that the gender PAY GAP was not solely a consequence of inequality in the pay of men and women for doing equal work in the same company. Other issues were also behind it, such as the tendency of women to opt for lower-paying jobs, to undertake more part-time work and to take career breaks due to child-rearing. Malta Today (vLNPG500, 2019-11, 137)

The Global Times is controlled by the Chinese Communist Party, and in Example 110 it positions COVID-19 as potentially frustrating closure of the gender pay gap, with the word *risk* suggesting that this would be undesirable. This is juxtaposed with the *quarter of men* who already feel it has *gone too far*. Significantly, the writer chose not to say the more obvious "more than three quarters of men" supported further closure of the pay gap, suggesting an ideological choice. The claim of *political correctness gone too far* can be viewed as an attempt to shut down discussion on a sensitive matter (Britannica [electronic resource], 2023).

Example 110 Efforts to close the gender pay gap risk stalling as countries recover from COVID-19, researchers said on Monday as a global poll showed almost a quarter of men thought the issue was " political correctness gone too far." Global Times (vLNPG500, 2021-03, 92)

These figures may refer to the International Women's Day Report (The Global Institute for Women's Leadership & Kings College London, 2021, p. 3) discussed in section 6.5.4

**5.4.3 Indifferentiation**

There are only nine examples of Indifferentiation in vLTPG500 and three in vLNPG500 (Table 55). At these levels Indifferentiation has little impact on the pay gap discussion, but is included here to allow comparison with the higher frequencies seen in the sexual violence datasets. Most of these pay gap examples are weak, but single examples from Twitter and Nexis are analysed below.

Themes within Indifferentiation						
Dataset (total positive Indifferentiation)	KWIC	Speaking out	Causes	Consequences	Denial	Abuse
vLTPG500 (9)	Pay Gap	3	0	2	0	0
vLNPG500 (3)	Pay Gap	5	0	0	0	0

*Table 55, Most common themes for positive Indifferentiation social actor types  
 Refer to footnote 76, page 135.*

An interesting perspective is provided by a transgender man Speaking Out in Example 111. The first sentence positions him as having personal experience, and his emotional statements *I can't stand the way I'm treated now, and it blows my mind & makes me sad* are powerful. Now presenting as male, he is treated with *more respect* suggesting a new perspective on a gendered culture. In calling out discrimination on behalf of those still burdened by it suggests solidarity, the hallmark of Indifferentiation.

Example 111 I have a unique perspective from being FTM<sup>82</sup>. I can't stand the way I'm treated now, with more respect given to me over a female, just because I now present as male; it blows my mind & makes me sad every time it happens.  
 The PAY GAP is just cruel and unjust. (vLTPG500, 2020-02, 257)

The New York Times, in Example 112, uses *champion* to suggest someone who would maintain the cause of, support and advocate (Oxford English Dictionary [electronic resource], n.d.-k), in this case for closing the gender pay gap. It alludes to the dynamic steps of *appointing a female tsar* and a *wider inclusion strategy*, again suggesting solidarity. The *However*, marks a change in tone, before warning that the solidarity needs to extend to the wider community, and translate to active steps if it is to deliver the desired outcome.

<sup>82</sup> FTM = female to male gender transition

Example 112 APPOINTING A FEMALE TSAR TO FIX THE PROBLEM Having an accountable champion for gender equality as part of a wider inclusion strategy is a sensible move. However, appointing someone purely because they are female, or appointing any figurehead without giving them full board support, won't help change your culture. It must be the responsibility of the whole leadership team, not just women, to take action on reducing the gender PAY GAP.

The New York Times (vLNPG500, 2019-04, 195)

### 5.4.4 Passivations

As discussed in Chapter 4, there is considerable overlap between Beneficialisation and Subjection within these datasets, and so these social actor types will be considered together in this section.

The Nexis Passivations are the most frequent positive social actor type making up 88% of that dataset, in contrast Passivations constitute only 61% of the Twitter dataset. Table 56 breaks Passivations into their main themes: Speaking out, Causes, and Consequences.

Themes within Pay Gap Passivations				
Dataset (total Passivations)	Passivation Type	Speaking out	Causes	Consequences
vLTPG500 (174)	Beneficialisation	82	63	20
vLNPG500 (371)	Beneficialisation	200	123	50
vLTPG500 (10)	Subjection	0	8	0
vLNPG500 (27)	Subjection	16	8	0

Table 56, Most common themes for positive Passivations social actor types  
Refer to footnote 76, page 135.

The contexts for these concordances are wide-ranging and include: gender, ethnicity; medicine; pocket money; tech industry; named universities; named companies; Hollywood; sport; and the women's soccer World Cup. The main themes will now be explored starting with Speaking out.

#### 5.4.4.1 Speaking out

Speaking out involves the use of communication verb phrases, and a common one is the verb *report*. The vLTPG500 contains many references to reporting by governmental bodies, universities, unions, professional regulatory bodies, and major employers. The language they contain is usually neutral, reflecting their corporate sources and accountable professional authors. For example the tweet in Example 113 which discusses the accountancy firm Ernst Young's (EY) reporting. The women employees, who are Suppressed, are also Beneficialised by an *improvement of 4.7%*. This tweet foregrounds EY's achievement while distancing the remaining gap of *32.4%*. The disadvantaged and invisible women are deduced from the nouns *gender* and *staff*.

Example 113 EY has reported a 32.4% gender PAY GAP among its staff, as of April 2020, an improvement of 4.7%. (vLTPG500, 2021-01, 87)

In Example 114 @spktralUK, a pay gap consultancy, uses *communicate culture and intent* to foreground the perception of doing the right thing, over the actual righting of a wrong. Use of the word *more* could imply a relative intention to be less unfair, rather than an absolute position of full diversity and zero pay gap. Women are Beneficialised by the *progress*, however the business phrasing suggests the focus is more on impressing potential clients or investors rather than the Suppressed female employees.

Example 114 At @spktralUK we are recommending the same to our clients. Gender PAY GAP reports are a very good way for you to communicate your culture and intent to make your environments more diverse and fairer as well as showing the progress you're making. (vLTPG500, 2021-02, 101)

During the Covid19 pandemic the UK government suspended Gender Pay Gap Reporting, its regulatory tool providing transparency. The writer of Example 115 questions the government's commitment with the question *if this govt is serious*, which suggests that currently the government is not seen as serious about closing the pay gap because it has suspended reporting. Women are also Beneficialised by the writer's exhortation to *sign the petition* in support of reinstatement.

Example 115 Sign the petition to get @trusslizz to Reinstate Gender PAY GAP Reporting - if this govt is serious about gender equality then it needs to make GPG reporting mandatory again. (vLTPG500, 2021-02, 396)

Example 116 adopts a more personal style in Speaking out. It is addressing earlier tweets by women (unavailable to this study) who are addressed as *sis*, short for sisters, suggesting a shared and personal alignment that cuts through demographic differences (Bouvier, 2022, p. 186). The intended audience *on this app* may be quite young considering the PS5 (PlayStation) reference. The fact that the writer needs to tell them *the pay gap still exists* suggests that many people remain unaware, or even deny the pay gap exists. The readers are also Beneficialised in a negative sense by the *1 an hour he makes over you*.

Example 116 Why are women on this app bragging about with they are buying men, sis the PAY GAP still exists. He can buy his own PS5 with the extra 1 an hour he makes over you... (vLTPG500, 2020-09, 93)

Women are also Beneficialised by the generalised *policy solutions* and the more specific *comprehensive childcare strategy* and *better data*, suggested by NERI<sup>83</sup> in Example 117, and again by @LisaWilsonNERI's *presentation* which promotes them.

---

<sup>83</sup> NERI = the Nevin Economic Research Institute, a trade union and research body.

Example 117 Fascinating presentation on the gender PAY GAP North and South from @LisaWilsonNERI @NERI\_research - policy solutions include comprehensive childcare strategy and better data - both of which @NIC ICTU has been calling for (vLTPG500, 2020-12, 267)

The reference to childcare may allude to the gendered burden of unpaid work that women try to balance against careers, that Nardone (2018, p. 181) found within German and Italian corpora, and Fineman (1995) describes as rooted in the definition of the traditional family.

The writer of Example 118 Speaks out to promote a podcast, in which they make associations between *spending power*, *advertising* and *stay-at-home mothers*. Here *women* and *mothers* are Beneficialised by a *ploy*, suggesting a “cunning scheme or manoeuvre” to gain a calculated advantage over them (Oxford English Dictionary [electronic resource], n.d.-y). Additionally, Scott Speaks out through her recently published book, *Double X Economy*, (2020) promoted in a podcast, creating layers of Speaking out.

Example 118 If you’ve not already listened to our @inspireajen’s chat with @ProfLindaScott - about women’s spending power, who advertising is really aimed at and whether the gender PAY GAP isn’t all a massive ploy to keep mothers at home - you should! (vLTPG500, 2020-08, 417)

The OECD report referenced in Example 119, highlights South Korea’s high pay gap. The use of *despite*, contrasts the injustice with South Korea’s *economic and technological advances*. This is evidenced by the low level of women in senior positions in business and government, suggesting that South Korean women are Beneficialised by a glass ceiling effect.

Example 119 Despite South Korea’s economic and technological advances, gender PAY GAP are among the highest in developed countries, according to an Organization for Economic Co-operation and Development (OECD) report. Women make up just 3.6% of the board members of business conglomerates. In terms of political representation, it is only 17% in the outgoing Parliament, which places the country at number 125 on the world list of the Inter-Parliamentary Union, not far from North Korea, which ranks 120<sup>th</sup>.  
Noticias Financieras English (vLNPG500, 2020-04, 453)

Example 120 ties together TSG’s reduced pay gap with the trialing of a *new flexible working day*, which would Beneficialise all employees. That they are *currently trialling*, leaves it unclear whether any benefit from this contributes to the reported figures. The communication verb *revealed* suggests a conscious act of unveiling (Oxford English Dictionary [electronic resource], n.d.-B) giving extra significance to the information provided.

Example 120 TSG, meanwhile, whose median PAY GAP fell from 36.7 per cent to 31.2 per cent year on year, revealed that it is currently trialling a new flexible working day policy aimed at helping staff achieve a healthy work/life balance.  
Computer Reseller News (vLNPG500, 2019-04, 49)

The Fine Gael speech verb *declared* is a formal assertion of their intention (Oxford English Dictionary [electronic resource], n.d.-l) to close the pay gap in Example 121. Fine Gael’s strong statement that the *pay gaps need to be eliminated next year* emphasises the necessity of doing so. In calling for it to be *eliminated*, rather than for instance *improving*

(Example 113) or *reducing* (Example 112) the pay gap, Fine Gael are considerably more optimistic than estimates from other sources. A search of the NexisPG corpus returns 59 concordances for the KWIC phrase YEARS TO CLOSE, and their values range from 26 years for Australian full time employees (Duncan & Holcombe, 2021) to 257 years for global figures (World Economic Forum, 2021). They may have meant equal pay which could be equalised more quickly.

Example 121 ‘Tackling gender pay gap a priority’ GENDER pay gaps need to be eliminated next year, a Fine Gael TD has declared.

The Sun (England) Ireland (vLNPG500, 2020-12, 20)

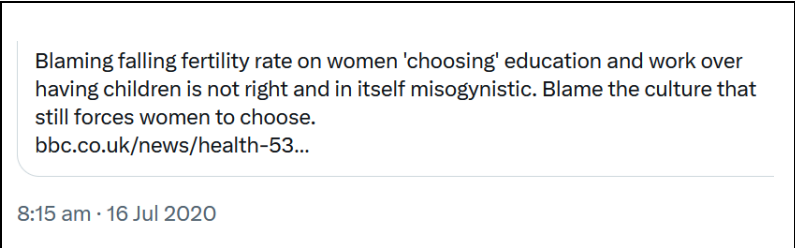
As a slightly unusual word choice a detour into a more general corpus was taken to clarify the general usage of *Tackling*. Within the enTenTen21 the top ten collocates of the verb *tackle* were: *challenges* (LogDice 7.89), *issue* (LogDice 7.89), *climate* (LogDice 7.67), *issues* (LogDice 7.23), *poverty* (LogDice 7.01), *head-on* (LogDice 6.95), *problems* (LogDice 6.88), *inequality* (LogDice 6.61), *inequalities* (LogDice 6.52). These words suggest that things needing to be tackled are typically serious, deeply embedded, and not easily overcome.

#### 5.4.4.2 Causes

The 63 Cause concordances within the positive Twitter Passivations are discussing (paraphrasing): awareness; culture; definition; transparency; and solutions.

Example 122 is a response to the tweet in Figure 18. Here the writer points out that arrangements for parental breaks often represent the father as a Suppressed social actor. The phrase *mum’s career takes a backseat*

Figure 18, Tweet “Blaming falling fertility on women ‘choosing’ education...”



Blaming falling fertility rate on women 'choosing' education and work over having children is not right and in itself misogynistic. Blame the culture that still forces women to choose.  
bbc.co.uk/news/health-53...

8:15 am · 16 Jul 2020

*because dad earns more* sums up the self-reinforcing negative relationship between *mum’s* decreasing earning potential and her increasing caring responsibilities. Cause and consequence are blurred, and become positioned as *practicality* in this pragmatic but inequitable situation.

Example 122 Also in the case of an equal partnership the decision doesn’t fall to one person! Both parties have a say. One issue I’ve seen is practicality. *mum’s* career takes a backseat because dad earns more. The PAY GAP still needs addressing. As does having to choose children OR career

<https://t.co/276VMlp8rL>

(vLTPG500, 2020-0, 182)

Example 123 develops the topic of motherhood as a cause of the pay gap, pointing to the contradictory tropes of women *can’t have it all* and *super mum*. The former suggests that mothers un-realistically want something that is not possible, while the later sets up an unattainable comic-strip fantasy standard against which they are judged. The verb *branded* invokes a sense of Subjection to a painful labelling, “typically one with negative connotations” (Oxford English Dictionary [electronic resource], n.d.-g) inverting the *super mum* epithet into something bitter. Here women are also negatively Beneficialised by un-

supportive working practices, and victim-blamed for causing the pay gap by working part time.

Example 123 Agreed. I also think there is something powerful and creative and enduring about parenting, which is rarely talked about. Instead we are ironically branded as super mum and simultaneously told we can't have it all, and the gender PAY GAP exists because we pt work. (vLTPG500, 2021-03, 370)

Unconscious selection bias in recruitment is offered as a contributory Cause of the pay gap in Example 124, but use of the relative expression *help reduce*, rather than the absolute *close* may be a recognition that it is not the only cause of bias in recruiting or promotion interviews.

Example 124 @h\_spencer of @Govt\_Equality at #inclusionMattersLdn :  
"Structured interviews can reduce bias and may help reduce the gender PAY GAP"  
" @MT\_editorial <https://t.co/WKs7iWAgHX> (vLTPG500, 2019-06, 316)

Example 125 offers Hanlon's razor<sup>84</sup> as underpinning the failure to eliminate the pay gap, describing the BBC's pay gap handling as just one in a list of other *shambles* due to one individual's *endless cock-ups* rather than by intention. The disadvantaged here are Beneficialised and Suppressed.

Example 125 @EddieHughes4WN @BBCBreaking Over 75's free licence fee U-turn, Proms PR disaster, gender PAY GAP shambles including Hall's disingenuous attempts to block public release of BBC salaries, EU referendum coverage bias, Sir Cliff Richard case outrage. His tenure's been characterised by endless cock-ups." (vLTPG500, 2020-08, 264)

The first sentence of Example 126 positions the topic as equal pay rather than a pay gap issue, but even this lesser requirement has *never been rectified*. Here, women are negatively Beneficialised by lower pay, bluntly attributed to the perception that men's work is *treated as superior*. Auspurg et al's (2017, pp. 202-204) lab-based study found men's work was valued more highly than women's, by both men and women as predicted by RET (Wagner & Berger, 1997).

Example 126 New Zealand eventually passed the 1960 Government Service Equal Pay Act and the 1972 Equal Pay Act, but men's work continued to be treated as superior to women's. The gender PAY GAP has never been rectified, according to a data release from Statistics New Zealand last week.

The Press (Christchurch, New Zealand) (vLNPG500, 2019-08, 102)

The two observations in Example 127, *fewer women in senior or higher-earning roles* and their greater need for *working part-time* are given as contributory factors in the pay gap. The writer acknowledges that *multiple, complex factors* operate here, and indeed the two factors could also be argued to be consequences of a glass ceiling and career segregation.

---

<sup>84</sup> Hanlon's razor states "Never attribute to malice that which is adequately explained by stupidity." (Wikipedia, n.d.-a)

Example 127 Sonya Lennon, stylist and founder of the WorkEqual campaign, said:  
“There are multiple, complex factors that contribute to the gender PAY GAP.  
These include fewer women in senior or higher-earning roles, and more women  
working part-time”. The Herald (Ireland) (vLNPG500, 2020-11, 296)

According to Example 128, conscious discrimination operates within tennis to influence the pay gap. In the Italian Open 2020, the women’s champion was denied parity with the male champion by €10. The writer points to outright discrimination being exercised *to keep the men on top*, with *however slightly* positioning this as petty. The punctuation and capitalisation emphasises the *10 EUROS*, a sum probably insufficient to buy lunch at the Italian Open. The writer suggests broad discontent saying *And fans were not happy*, illustrating this by quoting a fan to provide further support for their view. The writer clearly does not think the situation is funny, so the opening *Utterly hilarious* may be intended as sarcastic.

Example 128 Utterly hilarious that Rome, long an unequal prize money event, came SO CLOSE to giving equal prize money to both the finalists and champions today, but then decided to reduce the women’s prize money by 10 EUROS to keep the men on top, however slightly. And fans were not happy, with one writing: “People who made a PAY GAP in cents are misogynists. Period.”  
thesun.co.uk (vLNPG500, 2020-09, 212)

A study on discriminatory pay and prizes within the US Women’s World Cup team (Alkhamash, 2021, pp. 23, 27) found that social media users represented the governing bodies as responsible.

The Cause of direct Discrimination reported in Example 129, is represented as contributing 39% to the pay gap in 2019, and increasing. In contrast, Indirect cultural discrimination is suggested in Example 130, which calls for a list of changes, that by implication currently disadvantage women in Pakistan, including: leadership positions; involvement in key decisions; and a safe working environment. Taken together Pakistani women are Beneficialised by these systemic shortfalls.

Example 129 Researchers attributed gender discrimination to 39% of the gender PAY GAP, and noted its influence on the gap has increased since 2014.  
The Guardian (vLNPG500, 2019-08, 141)

Example 130 By acknowledging these nursing and midwifery leaders, WGH renews its call to WHO member nations to establish and support strategic leadership positions for nurses and midwives in health facilities, involve them in key decisions, support diversity in recruitment strategies, reduce the gender PAY GAP and ensure a safe work environment. Dawn (Pakistan) (vLNPG500, 2020-12, 168)

In recruitment it is common practice to ask applicants their current salary and use this as a basis for a job offer. Example 131 highlights the circularity created, locking already under-paid workers into ongoing under-pay. The phrase *pay history does not constitute a factor “other than sex”* acknowledges that the pay gap is gendered, and that women are negatively Beneficialised by their pay history.

Example 131 Finally, in 2020, the Ninth Circuit reaffirmed its prior decision and held that employers could not justify pay disparities by pointing to an employee's prior pay. It reasoned that pay history does not constitute a factor "other than sex" under the Equal Pay Act, and recognised that the use of a female applicant's prior pay to set her wages would perpetuate the gender-based PAY GAP. The Independent (vLNPG500, 2021-03, 399)

The West Australian, in Example 132, attributes the pay gap squarely to women's *unpaid caring and domestic work* compounded by unsupportive employment practices for mothers' returning to full time hours. The use of *disproportionate* alludes to unfairness, and *the motherhood penalty* positions motherhood specifically as something that is punished.

Example 132 For example, Australia's most female-dominated industry - health care - has had a gender PAY GAP increase in favour of men for the second year running. There's also the obvious - women's disproportionate share of unpaid caring and domestic work. Or the "motherhood penalty", where the gender pay gap widens after women take time off work to have children and find their earning capacity diminishes when they come back.

The West Australian (Perth) (vLNPG500, 2019-11, 287)

Both Example 132 and Example 133 refer to career segregation with respectively, the phrase *Australia's most female-dominated industry - health care*, and the more oblique *salary and career prospects associated with different jobs*. The unwritten message here is that traditional women's jobs are lower paid (Gould et al., 2016, p. 22), and that children are enculturated into different and gendered academic and career streams.

Example 133 "To reduce the gender PAY GAP, however, it is also important to make young people aware at an early age of the salary and career prospects associated with different jobs," he says. Die Welt (vLNPG500, 2021-03, 34)

Oracle spells out what career segregation means for its workforce in Example 134; with higher-paid jobs being held by STEM graduates, of which less than a third are women. This indicates that the awareness raising discussed in Example 133 needs to happen early and needs to be embedded before any subject specialisation.

Example 134 Their report read: "Most higher-paid jobs are filled by science, technology, engineering and maths graduates and Oracle's 25-30% female workforce is typical of the historic gender mix of these professional groups and graduate programs. This means there is a larger proportion of men in the higher paid jobs, which contributes significantly to the overall PAY GAP. "Oracle remains committed to reducing its gender pay gap by keeping diversity and inclusion at its core. Bracknell News (vLNPG500, 2019-04, 237)

A less discussed cause of the pay gap, described in Example 135, is its reciprocal relationship with domestic abuse. Controlling partners may prevent a woman from fully progressing her career, reducing her earning potential and ability to leave (Wagner & Berger, 1997, p. 20).

Example 135 All too often, she said, the gender PAY GAP is both a weapon and result of domestic abuse. “When we look at the victims who come into our shelter, they are very usually financially controlled, and financial abuse is one of the ways they experience domestic partner abuse,” Strzinek said.

The Enid News & Eagle (Enid, Okla.) (vLNPG500, 2020-12, 245)

The reciprocal action referred to in Example 135 refers to the abused individual as trapped in a lower paid role that prevents her moving out of an abusive home (Wagner & Berger, 1997, p. 20). The consequences of the pay gap are discussed further in section 5.4.4.3.

### **5.4.4.3 Consequences**

The consequences of the gender pay gap may manifest in all aspects of financial health. Example 136 focuses on US graduates having a *harder time repaying loans* associated with longer repayment periods, compounding their *student debt burden* and driving higher total costs. While financial hardship affects students of both sexes, *two-thirds* is held by women. According to the linked article, a further consequence of longer standing student debt is that “women often put off saving for retirement, buying a home or starting a business”.

Example 136 Women hold nearly two-thirds of the outstanding student debt in the United States -almost \$929 billion as of early-2019. Because of the gender PAY GAP, women have a harder time repaying loans, meaning women’s debt burden compounds as time goes on. <https://t.co/2vbyomwAA3>

(vLTPG500, 2021-03, 455)

Another consequence of the pay gap is the existence of Equal Pay Day, which gives advocates an opportunity to raise awareness and discuss progress, or lack of it. It is a measure of confusion around the terminology that it is called Equal Pay Day, although it refers to progress in closing the pay gap. In Example 137 *Happy Equal Pay Day* opens the tweet with an upbeat greeting. This creates a dramatic contrast with the second sentence in which The Salt Lake Tribune relates census data that ranks Utah *second-worst*.

Example 137 Happy Equal Pay Day: Utah has the nation’s second-worst gender PAY GAP, a new report says <https://t.co/XBzAH5SxQM> (vLTPG500, 2019-04, 373)

Equal Pay Day is referenced only twice in vLTPG500, but six times in vLNPG500. It could refer to one of two calendar days. One Equal Pay Day occurs before year end representing *the date on which women stop earning in a year, relative to men* referred to in Example 138. The other occurs after year end, explained in Example 139 as *the extra time it would take women, on average, to earn the same amount as men did that year*.

Example 138 According to the campaign, Equal Pay Day is the date on which women stop earning in a year, relative to men, because of the PAY GAP.

The Irish Times (vLNPG500, 2020-11, 110)

Example 139 Friday was 59 days since the end of the financial year - the extra time it would take women, on average, to earn the same amount as men did that year. The national gender PAY GAP remains at 14%, meaning men earn on average \$253.60 more than women per week.

Bendigo Advertiser vLNPG500, 2020-08, 53)

While statistics can be abstract and distancing from real world consequences, *40 per cent of single retired women live in poverty* foregrounds the human plight and the wider cost to society referenced in *need the support of taxpayers through welfare*.

Example 140 The gender PAY GAP means women are reaching retirement age with half as much super as men. Statistics show that Australian women retire with 47 per cent less super than men, only receive one third of the government's super tax concessions and as a result, 40 per cent of single retired women live in poverty. Thus, women are more likely to need the support of taxpayers through welfare.

Central Western Daily (vLNPG500, 2020-03, 78)

Sam Smethers, chief executive of The Fawcett Society, in Example 141, reframes the pay gap as a *productivity gap* highlighting the consequent loss to wider society of *women's skills and potential*. This perspective implicitly sets up closure of the pay gap as a win-win situation benefiting women and society.

Example 141 The PAY GAP represents a productivity gap and a waste of women's skills and potential. Too many women are trapped in low paid part-time work or locked out of nontraditional sectors while others experience pay or pregnancy discrimination. We need to address the underlying causes, one of which is the unequal sharing of unpaid care work.

Derby Telegraph (vLNPG500, 2019-11, 111)

### 5.4.5 Suppression

Suppression social actor types are four times as frequent in the vLTPG500 than the vLNPG500 (Table 57). These predominately have a Speaking out theme.

		Themes within Suppression				
Dataset (Total positive Suppression)	KWIC	Speaking out	Causes	Consequences	Humour	Denial
vLTPG500 (91)	Pay Gap	38	21	14	2	1
vLNPG500 (22)	Pay Gap	16	3	3	0	0

Table 57, The most common themes for positive Suppression social actor types  
 Refer to footnote 76, page 135.

**5.4.5.1 Speaking out**

Twitter is often used by the media, businesses and other bodies to promote their activities, such as that in Example 142. Here the European University Institute announces a forthcoming talk in which *@social\_charter* and *case-law of the European Committee of Social Rights* are foregrounded, but the disadvantaged individuals are not apparent.

Example 142 Join us this Thursday at the @FRWG\_EUI to talk about the gender PAY GAP in Europe. I will discuss equal pay under the @social\_charter and the recent case-law of the European Committee of Social Rights. (vLTPG500, 2021-03, 116)

In contrast, Example 143 starts with the emotional reaction *Bloody hell*, positioning the writer as speaking out in anger or surprise (Oxford English Dictionary [electronic resource], n.d.-f). The cause is the disclosure of a gendered difference in the size of donations to UK Labour party MPs (Figure 19). This funding is professional resourcing for running an office and so on, affecting an MP’s effectiveness and success. The writer seems to confuse these valid donations with illegal *bribes*. The woman receiving the highest donation is in 6<sup>th</sup> place, and there are only two women on the list of eight (25%). Following the 2019 general election 40% of UK MPs were female (Buchanan, 2025, pp. 10, 11) indicating that women are under-represented here.

Figure 19, Twitter graphic of MPs funded by Chinn donations

MP	Total	Payments
TOM WATSON	£66,475	13
KEIR STARMER	£50,000	1
DAVID LAMMY	£30,000	6
IVAN LEWIS	£30,000	3
OWEN SMITH	£27,000	4
ANGELA RAYNER	£25,000	1
IAN AUSTIN	£22,500	5
RUTH SMEETH	£10,500	3

Example 143 Bloody hell there’s even a gender PAY GAP for bribes. #Labour (vLTPG500, 2020-07, 168)

Similarly within Nexis, debate commonly focusses on the pay gap without referencing those impacted. For instance, in Example 144 Speaking out is enacted by *debate* and *disagreement*, but *gender reporting* and *ethnic pay levels* references do not specify which gender or ethnic groups are disadvantaged.

Example 144 Amid debate about the impact of gender reporting in reducing PAY GAP, there is similar disagreement as to whether forcing company disclosure on ethnic pay levels would have the desired effect.

Herald Scotland (vLNPG500, 2020-09, 167)

### 5.4.5.2 Causes

Both Suppression datasets refer to unpaid caring responsibilities as causal of the pay gap. In vLTPG500 this is often pitched as women's choices, while in vLNPG500 it more often frames unpaid care work and motherhood as a penalty or punishment. Discussion of causes such as bias in recruitment and promotion, lack of role models, and inadequate support for female progression are more commonly found in vLNPG500.

Both Example 145 and Example 146 position women's *choices* as being constrained. In the former, *Systemic biases* suggest issues that are "inherent in the system" (Oxford English Dictionary [electronic resource], n.d.-K) positioning the entire system as operating to constrain choices. Example 145 uses *from which it's hard...* to acknowledge the considerable effort needed to overcome the constraints. In Example 146 *not made in a vacuum* is a figurative reference (Oxford English Dictionary [electronic resource], n.d.-N) to these same unseen influences. While *Occupation and industry* are positioned as a *big part of the pay gap story*, *family responsibilities* and *sexism* are also listed as *factors* in Example 146, suggesting very wide issues.

Example 145 Choices do indeed play into the gender PAY GAP. But many of them are based on systemic biases from which it's hard, for people of any gender, to break free. (vLTPG500, 2019-04, 51)

Example 146 Occupation and industry choices are not made in a vacuum. They're a big part of the PAY GAP story - as of course are factors like family responsibilities and sexism. (vLTPG500, 2019-04, 496)

Prizes and salaries in sport frequently made media headlines around the 2023 Football World Cup and the 2024 Olympic Games, covering for instance:

- the Fédération Internationale de Football Association (FIFA) agreement to equal prize money in 2023 (Translated by Content Engine, 2023);
- the International Cricket Council (ICC) following suit in 2024 ("ICC enters era of equal prize money for men and women," 2024); and
- the absence of any women from Sportico's 2024 top 100 highest paid athletes (Badenhausen, 2024).

All of the social actors are Suppressed in Example 147, and the word *sexism* could refer to discrimination against either women or men. However, extratextual knowledge strongly suggests that it is women experiencing lower pay and prizes. This tweet is interesting for its suggestion that sexism against role models in competitive sport could *breed* discrimination beyond the sport. This is reminiscent of Alkhamash's (2021, pp. 29, 31) finding that US women's football players, who had a high profile on news and social media at the time, had become *cultural artefacts* in calls the change. A tweet from her study referred to the expression "Girls will be girls", a reversal of the more usual "boys will be boys", as being used

to challenge gender norms and given authority by “the importance of the Women’s World Cup win and highlighting the gender politics of women in sport.”

Example 147 asking why, in 2019, sexism is still so prevalent in tennis , and discussing how sexism in high profile careers breeds sexism in every day life  
(TwitterPG, 2019-05, 307)

Pay gap reporting is mandatory in more than half OECD countries including Ireland (2023, p. 19). The Irish Times, in Example 148, complains about the suspension of gender pay gap reporting (during the Covid19 pandemic) when it *has not been demonstrated that ... measurable progress in reducing the gender pay gap has been fulfilled. The committee carries a weight of authority to the pronouncement that closing the gap is a moral obligation.*

Example 148 The committee considers that the collection of reliable and standardised data is indispensable to the formulation of rational policy that would aim at reducing the gender PAY GAP. In the absence of such information . . . it has not been demonstrated that the obligation to achieve measurable progress in reducing the gender pay gap has been fulfilled.  
The Irish Times (vLNPG500, 2020-06, 432)

### **5.4.5.3 Consequences**

Fourteen of the vLTPG500 and three vLNPG500 Suppression concordances refer to the Consequences theme. Here again, there is some blurring between Causes and Consequences.

Example 149 is a direct statement regarding the unequal assignment of *benefit*. Systemic issues are stated as deliberate and causal in *institutions have been designed* with the consequence that this has been done in order to *benefit certain people*. This suggests Christie’s schoolchild disadvantaged by a system designed by another demographic, for their own benefit, discussed in section 2.5.1.1. The use of *certain* is interesting, emphasising that the writer knows who is involved, but has chosen not to specify the beneficiaries (Oxford English Dictionary [electronic resource], n.d.-j).

Example 149 Totally agree and brilliant points you made @rebgross about gender PAY GAP and the way institutions have been designed to benefit certain people @DrFBE  
(vLTPG500, 2019-09, 60)

Example 150 vividly expresses the consequences of the CEO pay gap that is *crippling solid people*. Increasing the minimum wage, associated here with the strong modal verb *will*, enable *solid people* to buy the things they need, also benefiting the traders they buy from, which consequently *will grow the economy significantly*. Women, suppressed from this tweet, are likely to occupy the minimum wage *solid people* position than be CEOs (McKinsey & Company & LeanIn.org, 2020, p. 10), and would benefit disproportionately.

Example 150 Raising the minimum wage to 15/hr will grow the economy significantly. The PAY GAP is crippling solid people at the base consumer level. 15/hr is a first step. CEO position salary can be cut 12% to cover this.

(vLTPG500, 2021-01, 150)

Nexis concordances also refer to the economic consequences for industry resulting from discriminatory practices, and represent the pay gap as *holding back future growth*, discussed in Example 151. The disadvantaged are suppressed here, with only the distanced and indirect references to *unpaid internships*, *workforces* and *lack of diversity* pointing to, but not identifying the missing.

Example 151 A major new report into Scotland’s film and TV sectors has raised concerns about the “widespread” use of unpaid internships, a lack of diversity across its workforces and a gender PAY GAP holding back future growth.

Scotsman (vLNPG500, 2019-09, 140)

### 5.5 Negative representations within pay gap datasets

This section explores the 116 (23%) negative concordances within vLTPG500 dataset. The frequencies of the main social actor types within the vLTPG500 negative concordances are repeated here as Table 58 for convenience. The most frequent social actor is Suppression (42), followed by Appraisal and Beneficialisation (19), and Activation (6). There are no negative concordances in the vLNPG500 dataset.

Dataset	KWIC	Activation Frequency (% of negative)	Appraisal Frequency (% of negative)	Beneficialisation Frequency (% of negative)	Subjection Frequency (% of negative)	Suppression Frequency (% of negative)
vLTPG500 (105)	Pay Gap	6 (7%)	19 (18%)	19 (18%)	4 (4%)	42 (40%)
vLNPG500 (0)	Pay Gap	0	0	0	0	0

Table 58, Frequency of negative representations of social actor types  
Refer to footnote 76, page 135.

Many concordances contained multiple themes, but for simplicity the two strongest themes are given in Table 59.

Theme	Activation Total 8	Appraisalment Total 23	Beneficialisation Total 19	Subjection Total 4	Suppression Total 42	Totals
Speaking out	1	1	1	1	1	5
Causes	4	3	5	0	6	18
Consequences	0	0	0	0	1	1
Denial	5	17	12	0	28	62
Abuse	2	3	4	2	5	16
Definitions	1	5	4	1	11	22
Humour/sarcasm	0	1	1	0	4	6
Other	5	11	8	0	16	40
Total	18	41	35	4	72	170

Table 59, Frequency of negative themes within the vLTPG500  
Refer to footnote 76, page 135.

These themes are discussed in the following sections.

### 5.5.1 Activation

The Activation concordances revealed five main themes, dominated by Denial (5) and Causes (4) and these will now be discussed.

#### 5.5.1.1 Denial

The most frequent theme within Activation is Denial with five concordances. It manifests as either denying the existence of a pay gap, or conceding that there is a gap but it is justified, usually due to women's choices. Here, the pay gap is defined as: *an absolute myth; false narrative; different career choices; and BS*. Denial within Activation concordances commonly co-occurs with Appraisalment, and where the Appraisalment feels more dominant these concordances are discussed in section 5.5.2.

Activation, in Example 152, is attributed to women who intransitively *fall pregnant*, suggesting a spontaneous event that Suppresses the father entirely. The phrase *focus their attention elsewhere* sounds like a euphemism for the burden of unpaid care work, reshaped here as a conscious action suggestive of victim-blaming. The final sentence is a passive-aggressive barbed compliment using women's independence as justification for condoning an unjust situation.

Example 152 as men. Women between the ages of 20 - 35 tend to fall pregnant / have children therefore they focus their attention elsewhere. The PAY GAP is an absolute myth. Why aren't "we" saying anything? Because you don't need our help and are strong independent women. (vLTPG500, 2019-09, 207)

Confusion of definitions is rife in pay gap discussions, and Example 153 is an illustration of this as the writer uses the word *pay gap* to refer to unequal pay for similar work. The writer's logic here is false; while she may agentically *make exactly the same as a man*, that doesn't

mean that the pay gap doesn't exist. Additionally, it is also possible that the writer is in a low-paid career segregated role.

Example 153 @Madeliney02 @ourhealthca @GavinNewsom BS on this PAY GAP. I make exactly the same as a man doing my job with the same amount of experience. (vLTPG500, 2019-09, 45)

The tone of Example 154 is unusually aggressive. The writer states their disinterest in not wanting to *read political books*, however they do want to publicly state their opinions. The second sentence follows on with *I act based on how you act* suggesting a reactionary response rather than discussion to gain understanding. They then ask *what the fuck do you people want*. *Modern feminists* here are Activated by *fighting*, however Denial is achieved by taking the post-feminist position, that the issue has already been *fought for and won*.

Example 154 I don't even read political books nor am I wanting to. I act based on how you act. And based on how modern feminists act, what the fuck do you people want lmao. PAY GAP - non existent, education - same. You're fighting for shit people already fought for and won. (vLTPG500, 2020-04, 89)

### 5.5.1.2 Causes

The verb phrases attributed to social actors within the Causes theme are: *use[ing] false narratives; marching and getting mad; make exactly the same as a man; [make] different career choices; fall pregnant/have children; bitch and moan; and act*. The individuals raising pay gap issues are represented as: *some ignorant female; vulgar, gross fools; morons; and haters*. All four Causes themed concordances also demonstrate victim-blaming.

In Example 155 the Activation uses *they bitch and moan*. The former suggests a "to grumble" or "complain" (Oxford English Dictionary [electronic resource], n.d.-d). The latter suggests grumbling, "typically about something trivial" (Oxford English Dictionary [electronic resource], n.d.-u). Together this conjures an impression of the pay gap as not serious enough to warrant complaint, and those calling it out are then victim-blamed for *act[ing] like vulgar, gross fools*. *Big bucks on endorsements* suggests that the topic is sport related.

Example 155 @marklevinshow Yes, morons. For one thing, they bitch and moan about the PAY GAP but then they act like vulgar, gross fools, thus reducing the likelihood that they'll make the big bucks on endorsements, etc. (vLTPG500, 2019-07, 230)

Dual coding of Activation with Appraisal occurs in Example 156 where the social actors are positioned as women who *don't marry down*, suggesting they prioritise wealth in a prospective partner. The writer positions women as being contrary; wanting to marry a higher paid man, but also wanting to earn equal pay. However, because the pay gap is defined by averages the average woman necessarily earns less than the average man, hence women are likely to earn less than their partner. Additionally, there is evidence that men may be more comfortable as the main breadwinner. Syrda's (2020, p. 976) survey of 6000 households found that men were least psychologically distressed when their female partners earned 40% of household income, and stress increased with women's income above that percentage.

Example 156 @LisaBritton (Identifying as a turnip for this post) I love hearing about PAY GAPS with laws that specifically prohibit paying people differently for the same work based on gender. The argument FOR a pay gap ... Women don't marry down. They want a guy who makes more money than they do.

(vLTPG500, 2020-02, 312)

### 5.5.2 Appraisalment

Within the 19 negative Appraisalment concordances the dominant theme is Denial (17) based on the pay gap being: *not a thing; BS; a lie; mythological; isn't entirely real; doesn't exist in America; debunked; false; not a pay gap; crap; nothing of substance; nonsense; a myth; stupid... thing*; and driven by *selfishness and more frivolous types of aspiration*. Women in these concordances are appraised as: [having the] *ability to just blame everyone but themselves, mediocre; having a gap between their ears, stupid*; and [you can't] *Show me a top female streamer that is as better at gaming*.

Example 157 positions the pay gap specifically not as a *myth*, but a *lie*, emphasising an "intent to deceive" (Oxford English Dictionary [electronic resource], n.d.-r), and perpetrated with the purpose of *manipulat[ing] other women*. A reader who might consider speaking out about the pay gap might be deterred by potentially being evaluated as part of the problem and harming other women. This example was also coded as a Cause.

Example 157 The Gender pay gap is not a myth. It's a Lie. Famous women (and men) say it every year to emotionally manipulate other women.

#GenderPayGap

(vLTPG500, 2019-09, 46)

The writer of Example 158 evaluates those who advocate for the existence of the pay gap as blaming *everyone but themselves*, a form of victim-blaming, and implying that those speaking out are the problem. Here a female employee is evaluated as *at best mediocre*, implying that she was generally below average, reducing her authority to *lecture* the writer. The word *lecture* Activates the female manager as speaking out "with some severity, or at length" (Oxford English Dictionary [electronic resource], n.d.-q).

Example 158 Lol the irony. That sucks, it's really their ability to just blame everyone but themselves is what amazes me. I have been similarly lectured about the PAY GAP by a female employee who was managing a team of six. While she was at best a mediocre employee.

(vLTPG500, 2020-08, 349)

Interestingly, the pay gap in Example 159 is assessed as being like feminism in having *too much equity baggage*. In stating that *too much equity, like the stupid pay gap thing*, is unreasonable, the text resembles the *political correctness gone to far* assertion, discussed in section 5.4.2. *Too much equity* seems to suggest an acknowledgment of unfair advantage, directly contradicting that the writer supports equality. The writer seems to use *humanist* to mean a kind of gender-blindness that fails to recognise systemic biases operating to maintain the status quo, rather than the spiritual position.

Example 159 @rosemcgowan I agree with that, and I also support equality, but feminism as it stands now doesn't represent that. it has too much equity baggage like the stupid PAY GAP thing. I'd sooner call myself a humanist.  
(vLTPG500, 2020-08, 494)

A Political tweet attacking both Joe Biden and the pay gap is seen in Example 160. The writer answers his<sup>85</sup> own equal pay question with an emphatic *No. They aren't*. Career segregation is positioned here as Latina women's choice, and consequently the pay gap is *nonsense*. The accusation of *pushing* the pay gap to create a job for his son, solving a *debunked* and therefore non-existent problem, judges Biden to be corrupt.

Example 160 @JoeBiden Are Latina women doing the same jobs? No. They aren't. This gender PAY GAP nonsense has been debunked countless times yet you still push it. Are you trying to get Hunter a job with the ACLU or something?  
(vLTPG500, 2019-11, 174)

Online gaming is notoriously hostile to women gamers (Marsh S., 2023), and in Example 161 the existence of any good female streamers is questioned with the *show me* phrase. The use of *This pay gap shit gets old* indicates that the writer is bored and impatient with calls for equal payment and prizes. Positioning female streamers as having weaker gaming skills also gives this example a Causes theme.

Example 161 This PAY GAP shit gets old My god. Show me a top female streamer that is as better at gaming then any of the top streamers...  
(vLTPG500, 2020-02, 477)

### 5.5.3 Passivations

The 23 Passivation social actor types are predominantly Beneficialisations (19) and will be discussed here together with the Subjections (4). Table 60 provides a breakdown of themes within the Passivation social actor type.

---

<sup>85</sup> The tweet meta data gives the user\_screen\_name as including a male title suggesting a man, and the user\_description includes "Marine by God!" which increases the probability of the writer being a man.

Themes within Pay Gap Passivations						
Dataset	Passivation Type	Denial	Definition	Causes	Abuse	Speaking out
vLTPG500 (19)	Beneficialisation	13	4	5	4	1
vLNPG500 (0)	Beneficialisation	0	0	0	0	0
vLTPG500 (4)	Subjection	0	1	0	2	1
vLNPG500 (0)	Subjection	0	0	0	0	0

Table 60, Most common themes for negative Passivation social actor types from the vLNPG500 and vLTPG500 datasets  
Refer to footnote 76, page 135.

Concordances with Denial, Definition, Causes and Abuse themes are discussed in the following sections. The Speaking out theme, with only single occurrences, did not reach the threshold for inclusion.

### 5.5.3.1 Denial

Denials often demonstrate confusion around the distinction between the pay gap and equal pay. For instance, the writer of Example 162 calls on the anonymous authority of *many economists and employers* who have *debunked* the pay gap. This tweet argues using the term *pay gap*, but by disputing the use of averages the writer also dismisses the definition of a pay gap. The *variants* may be a mis-spelling of variables, such as roles or hours worked, which are often used by deniers to justify lower pay, but are seen by proponents as contributing to the causes.

Example 162 The gender pay gap has been debunked by many economists and employers, were there is discrimination it should be challenged but the gender PAY GAP is a broad look at women working and getting a average, it doesn't take in a multitude of variants that could lead to lesser pay (vLTPG500, 2019-04, 22)

The pay gap is denied as *more feminist nonsense* in Example 163 where the writer misses the point that the pay gap by definition applies to the entire work force. Looking at equal pay by grade shows that, for mandated minimum wage roles the pay gap is non-existent because the cap creates a floor. At the highest levels of pay the gap is greatest, and this extends to distributions of bonuses and company shares (Cheng et al., 2018, p. 147). This tweet is accompanied by the visually powerful, with bright colours and capitalisation, graphic in Figure 20.

Example 163 I don't think it is legal for any job that falls under federal wage an hour laws to discriminate based on gender. Salaried and executive jobs are different. I also think the PAY GAP is more feminist nonsense.  
<https://t.co/SH7B8eNvlu> (vLTPG500, 2019-06, 332)

The graphic in Figure 20 acknowledges the pay gap, and seeks to justify it by pointing to *differences in occupations, positions, education, job tenure and hours worked*. However, these differences often map quite closely to male and female populations, for example hospitality workers are overwhelmingly female suggesting wider social issues such as career segregation and the glass ceiling.

Figure 20, Tweet "Fallacy of Understated Evidence" graphic

**THE FALLACY OF UNDERSTATED EVIDENCE**  
@otms.razor

**General fact: Women earn 77 cents for every dollar a man earns.**

**More specific fact: The 23-cent gender pay gap is simply the difference between the average earnings of all men and women working full-time. It does not account for differences in occupations, positions, education, job tenure or hours worked per week.**

**When you state a general fact to support your argument, but leave out the more specific facts that would contradict your argument.**

8:25 pm · 30 May 2019

### **5.5.3.2 Causes**

The concordances in this section were coded with the Causes theme, however the complex nature of the topic often means that the boundaries of Cause and Consequence are not clear cut.

Women cannot compete in the Association of Tennis Professionals (ATP) matches, but have to play under the different Women's Tennis Association rules. In Example 164 women are Beneficialised by being paid *exactly the same*. However, women's tennis currently has *less audience* (and less advertising) resulting from shorter match times contributing to women's game being negatively Beneficialised as 'less serious' (flyflayer, 2022). While the pay gap has been eliminated at Grand Slams, the *pay gap* in the last sentence, refers to the much larger number of lower level single-sex events.

Example 164 @BibianeSchoofs @FitLine\_Keli @RafaelNadal @rogerfederer  
Women get paid exactly the same at Grand Slams even if they play 3 setters.  
And they already are unified. The WTA brings in less audience than the ATP and  
in consequence generates less money. Doesn't that justify the 'PAY GAP' ??  
(vLTPG500, 2020-04, 33)

Alkahlammash (2021) drew attention to greater media attention to the men's game in US soccer, and gendered bias in naming that disadvantages female players.

Example 165 is a US political tweet attacking *the socialist way*. The writer here is discussing equal pay for similar jobs, pointing out that the *jobs are not identical*. This tweet uses the interesting phrase *White men might have more important jobs!* to negatively Beneficialise both women and ethnic groups with less important jobs. The disadvantaged may wonder whether those jobs became *important* because the incumbents are white men, because they are well paid, or perhaps because there are barriers to other demographics accessing these jobs. The writer justifies the status quo as *Just the way it is*, and follows through with *you are not bringing that up* which seems to end that line of discussion abruptly. The penultimate sentence demonstrates a colour-blind position, representing *the socialist way* as racist, and failing to recognise that systemic discrimination operates against less privileged demographics.

Example 165 Did you ever think maybe these jobs are not identical? White men  
might have more important jobs? Just the way it is and you are not bringing that  
up as the reason for the PAY GAP. Seems the socialist way is to take from the  
white people and give to everyone else, right??? Respond  
(vLTPG500, 2020-02, 180)

### **5.5.3.3 Abuse**

Several of the tweets discussed previously have a second theme of Abuse, one coded primarily as Abuse is discussed here.

Example 166 is patronising and insulting. *Sly* implies someone is deceitful or underhand (Oxford English Dictionary [electronic resource], n.d.-H) and this is compounded by the suggestion of a lack of *courage*. The *What is so difficult about this?* implies the individual is stupid in failing to understand *this* non-difficult thing. The final sentence states that *if*

women learnt the art of fact-based and rational confrontation, suggesting that they cannot currently do this, but if they could then it would help *erase the gender pay gap*.

Example 166 @can\_i\_be\_VIP If you have the courage to sly tweet, have the COURAGE to tag the person involved? What is so difficult about this? I promise if women learnt the art of fact-based and rational confrontation, we could do much to erase the gender PAY GAP. (vLTPG500, 2020-08, 64)

Rationality and Rational Legitimation are discussed in Chapter 6, where some examples are associated with equal pay arguments.

### 5.5.3.4 Definitions

Much of the discussion in these datasets appears to be at cross-purposes, with the pay gap being confused with equal pay, often referred to using descriptors like those seen in Example 167: *same job, same hours, same years of service*. Here, the pay gap is dismissed as a *myth*, and the writer alludes to greater injustices in *other countries* to invalidate the inequity discussed here. While other injustices may be greater, they do not belittle or alter the situation under discussion. The phrase *That's all you've got* is probably missing a question mark, and according to the Urban Dictionary (n.d.-a) the derogatory phrase is used in response to an opponent's greatest effort, feigning surprise at its inadequacy, and invoking a tone of self satisfaction intended to demoralise.

Example 167 Bullshit...show data. Anecdotes are not evidence of widespread pay gaps for same job, same hours, same years of service. PAY GAP myth! That's all you've got. I would hardly call this oppression of women. Go to some other countries if you want to see oppression. (vLTPG500, 2019-10, 210)

### 5.5.4 Suppression

Within the 42 negative Suppression concordances, 67% (28) show the Denial theme, 26% (11) Definition, 14% (6) Causes, 12% (5) Abuse, and 10% (4) Humour (Table 61).

Themes within Suppression						
Dataset	KWIC	Denial	Definition	Causes	Abuse	Humour
vLTPG500 (42)	Pay Gap	28	11	6	5	4
vLNPG500 (0)	Pay Gap	0	0	0	0	0

Table 61, Most common themes for negative Suppression social actor types  
Refer to footnote 76, page 135.

The themes associated with Suppression social actors combine in complex ways, but the following sections address the three most dominant themes, starting with Denial.

#### 5.5.4.1 Denial

The four examples discussed here as Denial all Suppress the people that they are discussing, removing them from the text. Three use *gender* to suggest women, illustrated by Example 168, justifying the pay gap with the generalised *you get paid what you're worth*.

The pay gap is also reduced to *a thing* in Example 169... In this case no justification is offered, but the writer describes their view of the subject with *Grumble*.

Example 168 There's no such thing as a 'gender PAY GAP'. You get paid what you're worth. Get over it. (vLTPG500, 2019-06, 142)

Example 169 But the PAY GAP isn't a thing! (Grumble.) (vLTPG500, 2020-05, 321)

A denial is accomplished with six words in Example 170, which positions *Gender pay gap??* as a question, and then answers it with *unicorns Not existing*. This creates an equivalence between the pay gap and unicorns, which are a mythical entity popular with children. The use of *Gender* implies it is women who believe this myth, so suggesting child-like gullibility.

Example 170 Gender PAY GAP ?? unicorns Not existing (vLTPG500, 2019-11, 381)

A brief return to SketchEngine revealed that the phrase *does not exist* occurs 83 (26.23 pmw) times in the 24 month TwitterPG corpus, almost all stating that the *pay gap doesn't exist*. In contrast, within the NexisPG corpus there are only four (1.38 pmw) occurrences, and one of these refers to another topic.

#### **5.5.4.2 Definition**

The second most dominant theme within negative Suppression social actor types is Definitions which has 11 occurrences. Two themes can be identified: gender is a social construct (4); equal pay and the pay gap as synonyms (4).

Four concordances highlight the social construction of gender which may suggest that career potential should be considered when adopting a gender, as different levels of remuneration are to be accepted as part of each gender's construction. Example 171 illustrates the issue's political polarisation by addressing *liberals*, other concordances address the *left* (Example 197) or *progressives* (Example 196). Ironically advocates of closing the pay gap also point to the social construction of barriers that underpin it (Government Equalities Office, 2019).

Example 171 Anyone else find it funny that the same people who claim there is a 'gender' PAY GAP are also saying that you define your gender? If one is true, the other can't be.. Take your pick, liberals. (vLTPG500, 2019-07, 441)

A further four concordances use the terms *pay gap* and *equal pay* as synonymous. Example 172's closing phrase *doesn't take in a multitude of variants that could lead to lesser pay* describes how equal pay differs from the pay gap. *Debunked* collocates with the lemma PAY GAP with a LogDice of 6.93 in the vLTPG500 corpus.

Example 172 The gender pay gap has been debunked by many economists and employers, were there is discrimination it should be challenged but the gender PAY GAP is a broad look at women working and getting a average, it doesn't take in a multitude of variants that could lead to lesser pay (vLTPG500, 2019-04, 22)

The confusion over definitions within the debate is highlighted in Example 173, providing evidence that at least some people are aware of it. The capitalisation of NOT and EQUAL suggests an element of frustration or irritation. This was the only example found of someone querying a definition.

Example 173 Please can I have acknowledgement from you that the gender PAY GAP is NOT about unequal pay for EQUAL work? (vLTPG500, 2019-04, 214)

### **5.5.4.3 Causes**

Within the seven negative Suppression concordances the Causes put forward are: *job choice* ; *pregnancy*; *working trends*; *females take lower risk jobs*; *variables*; and *women work fewer hours*. Advocates of the pay gap would recognise these issues, but view them as constraints arising from societal factors, rather than being free choices.

Career segregation is represented as a choice in Example 174, but this does not address why traditional women's careers are less valued. It isn't clear whether *until we start paying teachers the same as engineers* is a call to do so, or a statement of impossibility. Phrased as an instruction, *Do some reporting* accuses the reporter, @PoppyHarlowCNN, of not doing her job properly.

Example 174 @PoppyHarlowCNN not until we start paying teachers the same as engineers will the PAY GAP disappear. The vast majority of the pay gap is driven by job choice, not discrimination. Do some reporting. (vLTPG500, 2019-04, 374)

The nature versus nurture dichotomy is used in Example 175, but here reworked as *biology* versus *misogyny*. However, biology, like nature, generally refers to fixed and accepted attributes, a nurture based cause could be closed by cultural change.

Example 175 Yah, hasn't the gender PAY GAP been disproved multiple times? Isn't the real culprit just pregnancies and work choice? I.E. Biology not misogyny (vLTPG500, 2020-11, 195)

## **5.6 Humour and sarcasm**

Contradictorily, the only examples of writers explicitly attempting to be humorous are within vLTPG500, the most negative dataset.

Example 176 provides a fantasy response to a company's gender pay gap report. A riot scene within an office is described, with women *constructing shanks from plastic cutlery*. Although fantasy, the verbs *congregated*, *constructing*, *communicating*, and *running* make this one of the most agentic Activation representations. The *communicating via aggressive post-it notes* suggests that this rebellion has not entirely displaced everyday office behaviours.

Example 176 Hour 3 after my work published the gender PAY GAP report; Women have congregated in the far end of the office constructing shanks from plastic cutlery and communicating via aggressive post- it notes. Deborah running the show. (vLTPG500\_500, 2019-04 , 240)

### **5.6.1.1 Use of case as mocking**

Ambient humour in a social media environment is often difficult to analyse because utterances are brief and may lack context, but within the specific online community a reference may be homophobic, or taken as given (Zappavigna, 2012b, p. 152).

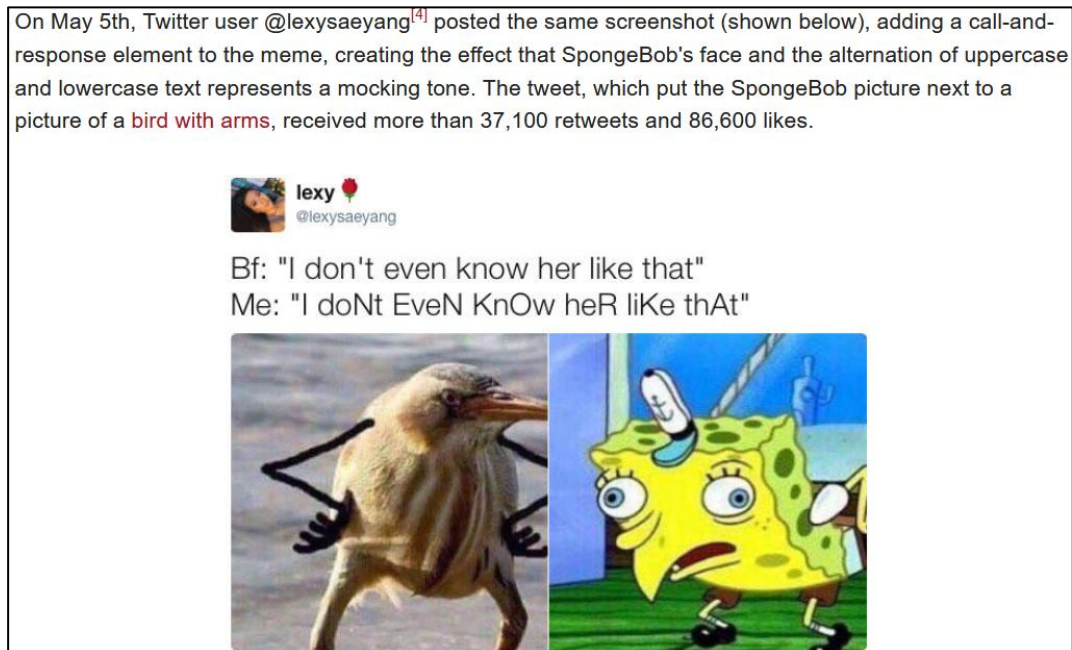
Humour is expressed in the fragment of a conversation in Example 177 which demonstrates the creative use of capitalisation to suggest a mocking response. The capitalisation of the last five words, signifying shouting, used to indicate that the speaker's agitation is being mocked.

Example 177 No one: Absolutely no one: Men in my comments: The pay gap  
dOeSN?t EXiSt IT? S LiTERALLY ILLEGAL TO DISCRIMINATE

(vLTPG500, 2019-07, 379)

Using random capitals as mockery, now common in social media posts, originated from the Mocking Spongebob meme (Walker A. & Matt, 2017). Figure 21 shows a snapshot of a tweet suggesting it's origin.

Figure 21, Tweet Mocking Spongebob meme



### 5.6.1.2 Creative use of punctuation

Using punctuation to create emoticons such as :) to indicate a smile, is long established, however greater creative use of punctuation is demonstrated in Example 178. Here, the writer refers to Utah, making a contrast between recent reporting of it having largest gender pay gap in the US (US Bureau of Labor Statistics, 2020, Table 3) while permitting gay marriage. Punctuation symbols are used to create a shrug  $\_ \_ (\_ \_ ) \_ \_ /$ , which suggests *disdain, indifference, or disclaiming responsibility* (Oxford English Dictionary [electronic resource], n.d.-F) although the upturned and skewed mouth may indicate a milder tone that is possibly bafflement, resignation or despair.

Example 178 So here we are in 2020, where gay marriage is legal and Utah is still  
one of the worst states for the gender PAY GAP.  $\_ \_ (\_ \_ ) \_ \_ /$

(vLTPG500, 2020-08, 151)

### 5.6.1.3 Sarcasm

*Income splitting*, referred to in Example 179, is a US reference to the allocation of income to family members to reduce overall tax liability (Criado-Perez, 2019, p. 258). This is most beneficial when one partner has unused tax allowance, as in the case of *stay-at-home moms*, and becomes less beneficial as the pay gap closes. The tone is sarcastic with *stay-at-home*

*moms are so last century* and having their children *growing up in daycare centres with strangers* being exaggeration, and one that mothers are not likely to support. This is an all or nothing argument used to create a dramatic effect, which misrepresents trained child care professionals as *strangers*. Child care outside the home is a necessity for single mothers and many working mothers.

Example 179 He also scrapped Income Splitting because 1) we all know the gender PAY GAP is BS, and 2) stay-at-home moms are so last century, surely children who would be better off growing up in crowded daycare centres with strangers.  
(vLTPG500, 2019-07, 505)

Figure 22, Tweet “Feminism page 1”

In contrast, another sarcastic tweet, writer B in Example 180 defends the existence of the pay gap in a way probably intended as humorous. Figure 22 is the seventh tweet in a thread between writers A and B, concerning feminists who *bitch about a pay gap*. Having set the scene with *I just learnt what feminism is*, B’s following *pay gap is page 1* suggests that the pay gap is something fundamental to feminism that would be stated on the first page of a book on the subject.



Example 180 Oh it was easy I just learnt what feminism is. Not derogatively using the word bitch or discrediting the PAY GAP is like page 1.  
(vLTPG500, 2020-10, 440)

The writer of Example 181 alludes to the trope that sex work is always an option for women, phrased here as *found a way to make money quick and easy* (Urban Dictionary, n.d.-b). The writer pretends not to understand why a reader would take offence with *y'all mad ??????*, emphasised by the multiple question marks. The writer is likely to be aware that their comment may be read as misogynistic, but if criticised they can claim it was “only a joke”.

Example 181 First of all why y’all complaining like a whole gender PAY GAP exist. Women found a way to make money quick and easy and y’all mad ?????? y’all need to start getting creative and stop being haters. (vLTPG500, 2020-04, 248)

## 5.7 Conclusion

The purpose of this chapter was to analyse the gender pay gap datasets and provide answers to research question 2.

The answer to RQ2A is yes and no. This is because vLTPG500 is more negative than all of the sexual violence datasets, but vLNPG500 is the least negative, having no negative concordances.

## Sexual violence and the gender pay gap

### 5 Prosody and the pay gap

Of all the datasets in this study vLTPG500 has the highest levels of neutral and negative (21%) concordances, and the lowest frequency (67%) of positive concordances. That all of the vLNPG500 concordances are positive, may reflect the contributions of professional journalists writing about large corporations with the power to sue them for libel. The news media reporting of pay gap stories tends to be unemotional reporting of the facts with few disputes, and interlocuters are usually large firms or government departments. This contrasts with the Nexis sexual violence datasets where the court cases that feature provide adversarial accounts where emotional insults, accusations, and denials can spill into the public domain, driving low frequencies of neutral and negative representations in the Nexis sexual violence datasets.

The answer to RQ2B, is that several themes arise for the pay gap: Activation, Appraisalment, Beneficialisation, Subjection, and Suppression, with the sub-themes: Speaking out, Causes, Consequences, Denials, Abuse, Definitions, and Humour/Sarcasm.

All of the themes mapping to social actor types present in the sexual violence datasets are also present for the pay gap, with the exception that Suppression is only associated with the pay gap. Suppression is the dominant theme within vLTPG500. The sexual violence datasets include the sub-themes of Blaming/shaming, within the pay gap corpora victim-blaming language was also identified but are less frequent and not appearing as collocates of *blame* and *shame*. Four new sub-themes emerge exclusively for vLTPG500 negative concordances: Abuse; Definition; Denial; and Humour/Sarcasm.

Differentiating between Cause and Consequence, often involving the same issues, has been a messy process. Many of the issues discussed are represented as either Cause or Consequence depending on the world view of the reader. It is possible that both are true, for example, women may choose traditional women's work and accept the associated lower pay for its flexibility, but constrained by social constructs such as the design of the traditional family and women's unpaid caring roles (Fineman, 1995). This results in women being career segregated into traditionally women's low paid roles.

The impact of unpaid care responsibilities is referenced throughout these English language pay gap corpora, and similarly Nardone (2018, p. 183) found them within German and Italian language. She concluded that "... pressing issues in society become recurrent topics in the language".

RQ2C focuses on the differences in themes between the pay gap and sexual violence datasets.

While many of the same themes occur for both pay gap and sexual violence datasets, they manifest in different ways. This is influenced by their fundamental differences, with the pay gap as a statistical measure, and sexual violence as a, generally, private act.

The most obvious difference is that of labelling. When they are referred to, those impacted by pay disparity are most often referred to using generic terms such as *women* or *employees*. Neither of these words foreground the sense of wrong in the way that *victim/survivor* does and this is suggestive of Fricker's (2007) hermeneutical injustice (section 2.5.9). This lack of a term analogous to *victim/survivor* contributes to those subject to the pay gap predominantly being positioned as Suppressed social actors, making up 28% of positive and 55% of its negative concordances within vLTPG500. In contrast, within vLNPG500 only 4% of positive

concordances contained Suppressed social actors. Chapter 7 continues this exploration of missing or silenced individuals.

The presence of the exclusively pay gap Definitions sub-theme provides evidence of the widespread confusion around the terms *pay gap* and *equal pay*. Within these corpora there is much less confusion about what sexual violence is. Equally the Denials pay gap theme disbelieves the existence of the pay gap as a phenomenon, often by incorrectly applying an equal pay argument to a pay gap situation. No one denies that sexual violence exists, although they might argue about who has experienced it, or whether a particular instance qualifies. The Denial theme is also exclusive to the pay gap and this is discussed further in Chapter 6, while the delegitimisation of Suppressed social actors is picked up in chapter 8.

## 6 Delegitimisation and the pay gap

### 6.1 Introduction

The purpose of this chapter is to address research question 3, which is repeated below for convenience.

RQ3, How are truthfulness and legitimacy attributed in the pay gap discourses, and how is this different from that seen in the sexual violence datasets? (Chapter 6)

This chapter analyses the full 24 month pay gap corpora, rather than the manually coded vL500 datasets of Chapters 4 and 5. Focus here shifts to representations of delegitimation, referring to van Leeuwen's Legitimation framework (2008, pp. 105-123). In *Discourse and Practice*, van Leeuwen (2008) largely discusses legitimation, and a search of the electronic version returns five instances of the word *delegitimation*, always alongside either *or* or *also*, suggesting that legitimation can be used in a positive or negative way (2008, p. 106). Within this study delegitimation is important, firstly because of its power to undermine the claims of those who are delegitimised, and secondly because patterns of delegitimation can be discriminatory when used against specific demographics. This chapter looks specifically at how delegitimation can be gendered in the pay gap corpora.

Ishani Maitra (2017, p. 287) uses the semantically close term *credibility* in her definition of silencing, saying that "credibility [is] unfairly deflated by racist/sexist prejudice". However, she goes on "authority may not be necessary for silencing (or subordination)" because if vilification is unrelenting, silencing may still be achieved even when meted out by a person who lacks authority. By this definition, those delegitimised by trolling or viral social media attacks may be viewed as muted or silenced.

The next section, 6.2, contrasts delegitimation in the pay gap corpora with that for sexual harassment, while the differences found between Twitter and Nexis pay gap texts are covered in section 6.3. The high level Veracity, Acceptability and Contempt themes are then explored in sections 6.4, 6.5, and 6.6 respectively. The implications of Chapter 6's analysis on RQ3 is explained in section 6.7.

Those missing from the pay gap discourse will be addressed in section 7.5.

### 6.2 Approaching the pay gap corpora

The first challenge was to find the best way into pay gap discourses. Viewing a SketchEngine WordSketch of *pay gap* suggested the adjective predicates of PAY GAP, in Table 62 columns one and three, and the "PAY GAP IS A ..." descriptors, in columns two and four.

The NexisPG corpus has fewer modifiers and much lower frequencies than TwitterPG. NexisPG has only two adjective predicates over the threshold for inclusion, *real* (LogDice 10.4) and *bad* (LogDice 10.1). No NexisPG "PAY GAP IS A ..." were above the threshold frequency of ten. Consequently, Nexis news articles do not feature heavily in the discussion of this chapter. The differences between the Nexis and Twitter corpora are discussed briefly in section 6.3

TwitterPG adjective predicate LogDice, (frequency)	TwitterPG "PAY GAP" IS A ... LogDice, (frequency)	NexisPG adjective predicate LogDice, (frequency)	NexisPG "PAY GAP" IS A ... LogDice, (frequency)
<i>real</i> 12.4, (418)	<i>myth</i> 12.9, (566)	<i>real</i> 10.4, (15)	<i>problem</i> 9.8, (7)
<i>bad</i> 10.0, (68)	<i>lie</i> 9.8, (49)	<i>bad</i> 10.1, (14)	<i>shocker</i> 9.5, (5)
<i>fake</i> 8.6, (21)	<i>problem</i> 9.8, (50)	<i>good</i> 9.1, (6)	
<i>false</i> 8.6, (21)	<i>bullshit</i> 9.2, (31)	<i>important</i> 8.6, (6)	
<i>wrong</i> 8.4, (20)	<i>hoax</i> 7.7, (11)		
<i>nonsense</i> 8.3, (17)	<i>joke</i> 7.4, (9)		
<i>ridiculous</i> 8.2, (17)	<i>fallacy</i> 6.8, (6)		
<i>important</i> 8.2, (18)	<i>farce</i> 6.8, (6)		
<i>fair</i> , 8.0, (15)			
<i>true</i> , 7.8, (13)			

Table 62, Evaluative adjective predicates and "pay gap" is a ... for the lemma PAY GAP  
WordSketch

Greyed cells indicate the terms falling below the frequency threshold of 10 which will not be discussed here.

*Real* (LogDice 12.4) and *myth* (LogDice 12.9) dominate Table 62 alongside a wide range of lower scoring evaluative modifiers, and these are grouped into three categories:

- Veracity, comprising *myth*, *real*, *lie*, *fake*, *false*, *true*, and *hoax* (section 6.4);
- Acceptability, comprising *bad*, *wrong*, *important*, *fair*, and *problem* (section 6.5); and
- Contempt, comprising *nonsense*, *ridiculous*, and *bullshit* (section 6.6).

Several of these words, particularly the Veracity group, create a markedness that sets up an implied unmarked alternative. For instance, if there is a *real* victim it implies there is also another un-*real* victim, who is therefore bogus. Kramer (2016, pp. 75, 76) describes how feedback loops between linguistic and social markedness may form, and become resistant to change.

The next section compares delegitimation in the Nexis and Twitter pay gap corpora.

### 6.3 Differences between the Nexis and Twitter pay gap corpora

Although both the Twitter and Nexis concordances often cover the same events, their writers have very different purposes and constraints, discussed previously in section 3.2.1.

SketchEngine WordSketches return a range of statistical terms as modifiers of PAY GAP. The LogDice scores for the two corpora show that Nexis is differentiated by its higher scoring statistical terms, such as *median* (LogDice 8.7), *mean* (8.0), and *average* (7.9), refer to Table 63.

Sexual violence and the gender pay gap  
 6 Delegitimisation and the pay gap

NexisPG includes the same words seen in TwitterPG but also has *analysis*. All of the LogDice scores within NexisPG, are higher than those for the TwitterPG.

The term *unadjusted* refers to the difference in average salaries for women and men within a given population, while *adjusted* figures have had some variables removed. The variables most often referred to include *jobs*, *experience*, *hours* or *danger*, hence the *adjustment* brings the definition into line with equal pay.

NexisPG freq (LogDice)	TwitterPG freq (LogDice)
<i>median 201 (8.7)</i>	<i>median 78 (5.3)</i>
<i>mean 118 (8.0)</i>	<i>average 72 (5.2)</i>
<i>average 124 (7.9)</i>	<i>mean 37 (4.3)</i>
<i>analysis 4 (7.6)</i>	<i>adjusted 20 (3.4)</i>
<i>unadjusted 34 (6.2)</i>	<i>unadjusted 14 (2.9)</i>
<i>adjusted 22 (5.6)</i>	

Table 63, Statistic related modifiers of PAY GAP for NexisPG and TwitterPG

The NexisPG texts are more likely to recognise the drivers of the gender pay gap, illustrated in Example 182, where the Chicago Herald differentiates and addresses both equal pay and the gender pay gap.

Example 182 But having little to no ADJUSTED or unadjusted gender pay gaps really sets them apart."Starbucks' parity shows not only that women get "equal pay for equal work" but also that they have achieved as many high-paying roles as men. Chicago Daily Herald (NexisPG, ADJUSTED, 2019-12)

In contrast, TwitterPG concordances often use *unadjusted* to position a claim as invalid, as in Example 183, which discounts the unadjusted figures as *nonsense*, and not *sane*. The tweet attacks the linked Forbes magazine article as being *equal pay day nonsense*. The writer here seeks to adjust away several causes of the pay gap including: field choice (*career segregation*), and *career breaks* (unpaid caring responsibilities). In positioning the pay gap as *nonsense*, the people disadvantaged by it are rendered invisible, or delegitimised. However, Example 183 does concede that there may be some *systemic discrimination*.

Example 183 Try this. Or any sane analysis using ADJUSTED pay gap for hours worked, degree and field choice, career breaks, etc. More like 5% unexplained or attributed to systemic discrimination. Not the 30% or that April equal pay day nonsense. <https://t.co/hE8FOUIXKz> (TwitterPG, ADJUSTED, 2020-02)

Career segregation is commonly discussed in these corpora, and often represented as women's choices. However, by definition segregation is not chosen, but imposed on a community by a more powerful entity (Oxford English Dictionary [electronic resource], n.d.-D). Throughout this chapter those whose arguments use either equal pay/adjusted figures or pay gap/unadjusted figures, often seem to be communicating at cross-purposes.

The next section moves on to explore the first of the three themes arising within delegitimation, Veracity.

## 6.4 Veracity

Veracity is a prominent theme in pay gap concordances that prioritises a focus on truth but potentially ignores power and privilege (Chouliaraki, 2024, p. 17). In Table 62 there were seven adjective predicates and “pay gap” is a ... terms relating to veracity with five clearly expressing untruth: *myth* (LogDice 12.9), *lie* (LogDice 9.8), *fake* (LogDice 8.6), *false* (LogDice 8.6), and *hoax* (LogDice 7.7). Only *real* (LogDice 12.4), and *true* (LogDice 7.8) purport to express truth rather than falsehood, however this turned out not to be the case because these were used in inversions, described in sections 6.4.2 and 6.4.7, leaving all Veracity concordances with negative prosody. The following sections explore some examples of Veracity concordances.

### 6.4.1 *Myth* - TwitterPG

Among the 566 occurrences of *myth* (LogDice 12.9) within TwitterPG, only 16 (3%) are part of a *not a myth* construct, with 12 (2%) caveated as *largely/mostly a myth*. Within NexisPG there were no *myth* collocates.

Aside from straightforward bald denials of the pay gap, three subthemes of *myth* emerged (paraphrasing): women’s choices; liars peddling myths to deliberately hurt other people; and if there was a pay gap, companies would only hire women. Discussion of these three themes follows in sections 6.4.1.1, 6.4.1.2, and 6.4.1.3.

#### 6.4.1.1 *Women’s choices, and introducing control*

Throughout the TwitterPG corpus there are many references to the pay gap as the natural result of women’s choices, illustrated in Example 184 using commonsense Rational Legitimation. The word *choices* attributes agency to *women*, along with the responsibility for securing a good salary. While agency is normally a positive attribute, here the possible restrictions on women’s choices are ignored, positioning women as responsible for their own disadvantage, effectively victim-blaming them.

Example 184 @SenGillibrand The PAY GAP IS A MYTH based in the different choices of men and women, as well as differences in experience and salary negotiation. When these are controlled for, the pay gap is about 1%.

(TwitterPG, PAY GAP IS A MYTH, 2019-04)

The term *controlled* (LogDice 2.5) seemed to carry an unexpected prosody that was seen widely across this study, so this prompted a detour to investigate further. The enTenTen21 corpus was used as a comparison of the use of *controlled* across the internet more generally. The search terms were CONTROLLED FOR and UNCONTROLLED, the results are shown in Table 64. The CONTROLLED FOR collocates in column 1 are all strongly scientific terms suggesting necessary scientific rigour. In contrast, the top 10 collocates of UNCONTROLLED include six chronic health conditions, and two referred to research: *open-label* (LogDice 5.71), and *intercurrent* (LogDice 5.81). *Airspace* (LogDice 6.59) refers to air-traffic risk, and *proliferation* (LogDice 6.39) refers to weapons and cancer. This suggests that CONTROLLED FOR has strong positive scientific prosody, while UNCONTROLLED suggests negative prosody involving chronic health conditions or safety-critical situations.

Sexual violence and the gender pay gap  
6 Delegitimisation and the pay gap

Collocates of CONTROLLED FOR (LogDice)	Collocates of UNCONTROLLED (LogDice)
Controlled for <i>confounders</i> (7.37)	Uncontrolled <i>hypertension</i> (7.91)
Controlled for <i>confounding</i> (6.92)	Uncontrolled <i>airspace</i> (6.59)
Controlled for <i>covariates</i> (6.19)	Uncontrolled <i>diabetes</i> (6.51)
Controlled for <i>thermostatically</i> (5.25)	Uncontrolled <i>asthma</i> (6.5)
Controlled for <i>socio-demographic</i> (4.93)	Uncontrolled <i>proliferation</i> (6.39)
Controlled for <i>confounds</i> (4.8)	Uncontrolled <i>seizures</i> (6.13)
Controlled for <i>multivariate</i> (4.78)	Uncontrolled <i>mellitus</i> (5.83)
Controlled for <i>analyses</i> (4.6)	Uncontrolled <i>intercurrent</i> (5.81)
Controlled for <i>regressions</i> (4.6)	Uncontrolled <i>open-label</i> (5.71)
Controlled for <i>endogeneity</i> (4.34)	Uncontrolled <i>and unregulated</i> (5.56)

Table 64, Top ten collocates of CONTROLLED FOR and UNCONTROLLED within the enTenTen21 corpus

Returning to TwitterPG, a WordSketch for CONTROL returns the prepositional phrase *controlled for* with a LogDice of 37.9 the highest score seen anywhere in this study. Further down the rankings were: *variables* (LogDice 10.88), *factors* (LogDice 9.71), *motherhood* (LogDice 8.97), *experience* (LogDice 6.67), *hours* (LogDice 6.57), *job* (LogDice 6.34), and *education* (LogDice 6.24). These elements are represented as commonsense, natural and reasonable justifications that a pay gap needs to CONTROL FOR, i.e. remove.

The writer of Example 185 has spotted the flaw in this argument, and uses *height* as an amusing analogy. Most women are shorter than most men, so if short people are less valuable, then most short people/women should be paid less than most men.

Example 185 Once you CONTROL FOR height, there is no gender pay gap. It is just a height pay gap. (TwitterPG, CONTROL FOR, 2020-01)

However, height is not generally factored into salaries, and furthermore height is fixed, while the variables listed in the text above as collocates of CONTROLLED FOR in Table 64 with LogDice 6.67 to LogDice 6.24 are social constructs, and therefore changeable.

### 6.4.1.2 Liars hurting other people

The second theme positions those calling out the pay gap as expressly motivated to hurt innocent others. A myth is a well-known but untrue story. Illustrated in Example 186, the *myth* is represented as having been *deliberately propagated* by *feminists* who are responsible for the damage it causes. The deliberate and large scale nature of the damage is emphasised by *manufacture unhappiness & resentment*. The writer also uses the flawed logic that, because in most western countries there is an Equal Pay Act (for example US 1963, UK 1970) equal pay is guaranteed. The reference to longstanding law adds a sense of Traditional Legitimation. Individuals reading this may be deterred from speaking out, afraid of being portrayed as: a delegitimised *feminist*; an *unhappy & resentful* and Passivated *woman*; or as a greedy person who wants to be *paid more for working less*. Here the only socially acceptable, but unspoken, alternative appears to be remaining silent.

Example 186 lol. The PAY GAP IS ANOTHER FEMINIST MYTH deliberately propagated to manufacture unhappiness & resentment. It has been illegal to pay a woman less than a man for the same job for 40 years. What the pay gap pretends is that women should be paid more for working less

(TwitterPG, PAY GAP IS A MYTH, 2020-07)

### 6.4.1.3 Companies would hire only women

The last of the three major themes for PAY GAP IS A MYTH is illustrated by Example 187. This popular theme argues that if women really did earn less than men, then *companies would only hire women*. While this is an equal pay argument, it quite frequently features in pay gap discussions. Commonsense appeal gives it Theoretical Rationalisation Legitimation as it would then follow that because companies do hire men, then the pay gap can't exist. This argument oversimplifies recruitment as being based exclusively on salary when this is not the case. In reality, when companies hire men rather than equally qualified women, pay men more than women in similar roles, or promote men over better qualified women, it is because they perceive men's work as more valuable (Auspurg et al., 2017, pp. 180, 202). Cooke (2024, p. 357) refers to an "ideal worker" who has total commitment to their job and no un-paid caring responsibilities. This results in all women, whether or not they have care responsibilities, being viewed by employers as potential carers, and less productive. Within a household, the partner who is paid less, is more likely to have to take on any additional unpaid caring activities, and then require more flexibility from employers to meet the demands placed upon them, making this belief self-fulfilling. When this partner is generally of one sex the issue becomes gendered.

Example 187 Gender PAY GAP IS A MYTH. If it were true, all companies would only hire women and cut their payroll by 28%. (TwitterPG, PAY GAP IS A MYTH, 2019-06)

*Myth* concordances call on Rationalisation to legitimise and add authority to their comments, resulting in the delegitimisation of those advocating for closure of the pay gap.

### 6.4.2 Real - TwitterPG

*Real* has the highest score (LogDice of 12.4) of any evaluative adjective predicate on Table 62 and carries more positive prosody than other Veracity modifiers. A manual count of the *real* concordances showed that 67% state that the pay gap is *real* and a further 4% position it as *very real*, refer to Table 65. Other uses include the speculative *if the pay gap was/were real* and *isn't/not real* featuring in 9% and 15% of concordances respectively.

Use of <i>real</i>	Frequency	Percentage (of the full 418 <i>real</i> concordances)
<i>is real</i>	279	67%
<i>if the pay gap was/were real</i>	39	9%
<i>is not/Isn't real</i>	61	15%
<i>very real</i>	18	4%

Table 65, Breakdown of adjective predicates of *real* within TwitterPG that exceed the frequency threshold of 10.

The phrase *if the pay gap was/were real* is always accompanied by purported evidence that the pay gap is not real, calling on Rational Legitimation. The present tense, *if the pay gap is real*, occurs in only seven of the 39. The conversation is always about a currently existing pay

gap, so it is interesting that *was* and *were* predominate, repositioning and distancing the pay gap into the past.

The 67% of *is real* adjective predicate concordances are most often part of a wider construct that inverts the meaning. A WordSketch Concordance search limited to *pay gap is real* as the KWIC in the TwitterPG corpus found 111 concordances (a subset of the 418 referred to in Table 64) and a manual coding of these, one of which shown in Example 188, showed that 43 were denying the pay gap.

Example 188 Oh my God, the pay gap crap again, yet nobody can truly verify that  
PAY GAP IS REAL outside of working hours etc (TwitterPG, 2019-05)

Within these *real* concordances four sub-themes emerge (paraphrasing): women's choices; companies would only hire women; biology or natural differences; and politics. The first two were also associated with the *pay gap myth*, while the latter pair appear exclusively with the *pay gap is real* concordances. These themes are discussed in the following sections.

#### **6.4.2.1 Women's choices**

In Example 189 *Bigfoot* refers to a mythical creature from US folklore used here to create a Moral Evaluation Mythopoesis analogy. Assuming a full stop between *Bigfoot* and *women*, this concordance represents the *opt in to lower-paying jobs* as women's choice, suggesting that given the choice of similar roles, women would choose the lower paying option. The writer goes on to point out that women *work less hours*, and *spend more time un- or under-employed*, positioning these as further choices, rather than reluctant compromises to accommodate unpaid caring work. The term *voluntarily* emphasises free choice "without external constraint" (Oxford English Dictionary, n.d.) which may not reflect reality for many women.

Example 189 the gender PAY GAP IS AS REAL as Bigfoot women tend to voluntarily opt  
in to lower-paying jobs (nurse > doctor; teacher > engineer), work less hours,  
spend more time un- or under-employed, etc.  
(TwitterPG, PAY GAP IS REAL, 2019-07)

Occasionally writers manage to counter delegitimation with humour. The tweet in Example 190 reframes the attraction that low paid roles hold for women.

Example 190 One explanation for the gender PAY GAP is that men gravitate towards  
higher paying jobs like doctor, engineer, CEO... while women tend to gravitate  
towards lower paying jobs like female doctor, female engineer and female CEO.  
(TwitterPG, PAY GAP, 2019-04)

#### **6.4.2.2 Biology, natural differences**

The use of *we* suggests that the writer of Example 191 is a woman, which would lend Personal Authority to this unusual biological argument that positions the *gift of bearing child birth* as compensation for unequal pay and presumably the pay gap. This reflects the traditional and conservative view that women should be content with motherhood and deprioritise careers. This argument stops short of explaining how this would work for single mothers and households needing two full-time wages, or why this would affect the salaries of childless women. The second argument mistakenly believes that the request is for pay

parity across different jobs, while the references to *hours worked and skills level differ* echo the variables that should be adjusted for, discussed in section 6.3.

Example 191 ???? This is an important issue that a lot of people are still struggling with. The GENDER PAY GAP IS NOT REAL . Women have the gift of bearing child birth, we can't expect to be remunerated the same as a man of [if] the hours worked and skills level differ. (TwitterPG, PAY GAP IS REAL, 2020-08)

More common among the *real* concordances is the positioning of social constructions as natural and fixed differences between men and women. Example 192 explains the impact of women's work on *the home front*, as natural and taken for granted, without explaining why women alone are responsible for this work. The writer also points out that conscious discrimination may not be involved, aligning with Auspurg's lab-based study (2017, p. 202). Another common argument refers to the perception that women, as an innate female trait, have weak negotiating skills. However, Cheng (2018, p. 146) found that women acting assertively (non-normatively) are punished in the workplace, but conversely, acting normatively and performing stereotypical feminine traits, they are viewed as weaker and made worse offers and then put under more pressure to accept them.

Example 192 Hope you know the gender PAY GAP IS NOT REAL ? Work places don't intentionally decide not to pay women as much as men. For one, women don't advance as much in their careers due to the home front. Also, it could be because the woman didn't negotiate her pay as hard as the man did. (TwitterPG, PAY GAP IS REAL, 2020-12)

Example 193 shows a different slant on natural differences. The writer states *Talent is universal* and the similar *Obtainable skills are universal*. Both of these assume a level playing field where all have equal opportunities, and talent and effort are the only influences on who rises to the top and gets the best paid positions. The tweet is given more weight by the Instrumental Legitimation of *teaching others that everyone has the power to be financially free in this country*. This phrase foregrounds the writer's altruistic act, but suggests a world view where low pay results from lack of talent or effort on the part of the individual. This argument ignores the societal barriers facing some demographics, while simultaneously justifying the position of those currently in powerful and well paid roles. This delegitimises those disadvantaged by the pay gap as lacking talent or effort, effectively victim-blaming them, which may deter them from speaking out.

Example 193 I love money. I don't love it over people and values... but I love money and teaching others that everyone has the power to be financially free in this country. The gender PAY GAP IS NOT REAL ..??????? Talent is universal. Obtainable skills are universal. (TwitterPG, PAY GAP IS REAL, 2019-09)

### **6.4.2.3 Politics**

Particularly within the US tweets, the pay gap is a very politically polarised issue, with support for closing the pay gap almost exclusively from Democrats (the political left), and denial of it from Republicans (the right). Palomino Manjón (2024, p. 119) found that Republicans viewed #MeToo and accusations of sexual violence as weaponised against them by Democrats and Example 194 suggests that thinking is extended to the pay gap. Example

194 illustrates delegitimation with *Susie* being indirectly accused of lying with the phrases *You know it's not real* and *pretend*. The use of *Actually* as the first word sets up the tweet as about to disagree with *Susie's* position (Oxford English Dictionary, n.d.-a).

Example 194 Actually *Susie* the PAY GAP IS NOT REAL . Democrats just use it for fund raising and as a political weapon. You know it's not real but like all democrats use it to pretend you stand against injustice. How about find something real to stand on. (TwitterPG, PAY GAP IS REAL, 2019-06)

#### **6.4.2.4 Companies would hire only women**

Previously discussed in relation to *myth* in section 6.4.1.3, the theme *Companies would hire only women* will not be repeated here. However, Example 195 has a slightly different twist, appearing to be a sarcastic response mocking a previous tweet. While *prejudiced* may have been intended sarcastically, there is evidence that unconscious bias by all parties, including women, does operate in the workplace (Auspurg et al., 2017, pp. 202, 203). The final sentence conflates legislation with compliance, which is analogous to saying that burglary is against the law, therefore burglary doesn't exist. The choice of *forcing diversity hires* positions equality of employment opportunities as being unwelcome and only undertaken to avoid breaking the law.

Example 195 Yeah, companies don't hire minorities or women because they're lazy and prejudiced. If the gender PAY GAP IS REAL, and companies only care about the bottom line, wouldn't they only staff women? Furthermore; There are laws in place forcing diversity hires already, you dunce. (TwitterPG, PAY GAP IS REAL, 2019-11)

The final *you dunce* represents the individual as being slow of learning or stupid, a delegitimation using an inversion of van Leeuwen's Expert Authority type. By implication this insult is also aimed at any others who believe that equal pay legislation has not eradicated the pay gap.

#### **6.4.3 Real - NexisPG**

In contrast to TwitterPG, NexisPG concordances use *real* to legitimise the pay gap. Of the 15 *real* adjective predicate concordances in the NexisPG corpus, 12 identify Democrats, with ten being direct quotes from President Biden. Eight of these reference the US women's soccer team on Equal Pay Day 2021.

Example 196 is fairly typical of these in that it uses the term pay gap to discuss equal pay. The President of the United States's endorsement that the *pay gap is real*, arguably gives the highest level of legitimation possible within the US. Furthermore, the epithet *soccer stars* represents Rapinoe and Purce as having authority as Role Models, while the title *Dr.* positions Jill Biden as an Expert authority through her professional status, alongside her social standing as *first lady*. All three of these women also claim Personal and Experiential legitimation as they recount personal experiences of unequal pay. Rapinoe describes her personal experience of being *devalued*, *disrespected* and *dismissed*, which she attributes to gender discrimination (Cooke, 2024, p. 358).

Example 196 Soccer stars Megan Rapinoe and Margaret Purce joined President Joe Biden and first lady Jill Biden to mark Equal Pay Day at the White House complex. 'It doesn't matter if you're an electrician, an accountant or part of the best d\*\*n soccer team in the world, the PAY GAP IS REAL,' the president said. Both Rapinoe and Dr. Jill Biden told anecdotes about being paid less than their male counterparts. 'I've been devalued, I've been disrespected and dismissed because I am a woman,' Rapinoe said

Dailymail.com (NexisPG, PAY GAP IS REAL, 2021-03)

In Example 197 only one of NexisPG's 15 concordances represents closure of the pay gap as (mildly) negative, being *largely misrepresented by the left*. The writer accuses the left of failing to effectively communicate the true meaning of the pay gap. While the statistical calculations find that at population level women earn less than men, the language used by the left leads *most people* to believe that there is an equal pay failure within specific jobs. This could arise from the common confusion seen in the data of this study between the terms *equal pay* and the *pay gap* by writers of all political positions.

Example 197 Although the gender PAY GAP IS REAL, it is largely misrepresented by the left. To me, the type of rhetoric from the left about the gender pay gap in America is mind-boggling. We constantly hear that women only earn 80 cents for every dollar earned by men. From this, most people infer that a man will be paid more than a woman for doing the same job and working the same number of hours. The Mirror: Fairfield University (NexisPG, PAY GAP IS REAL, 2019-10)

While equal pay remains an ongoing problem in some areas, most women who are disadvantaged are affected by pay gap issues such as lack of child care, career segregation, and/or a glass ceiling rather than equal pay failure.

#### 6.4.4 Lie - TwitterPG

WordSketch found no occurrences of the *pay gap is a lie* within NexisPG, in contrast to TwitterPG where the phrase had a LogDice score of 9.8. A manual check of the 49 TwitterPG concordances revealed that lies originated from: *feminists* (9), *politics* (5), *[the] west* (3), *liberals* (2), *Marxists* (1), *democrats* (1), *the insecure envious far left* (1), and *virtue signallers* (1).

All of the terms listed are associated with the political left, suggesting pay gap deniers are more aligned with the political right, which views individuals as responsible for their own outcomes, and supports traditional values such as the non-working mother. The alignment of denying tweets with populist and neoliberal politics is discussed more fully in Chapter 7.

The writer of Example 198 states that *My wife my mother my sister* have equal pay, establishing Experiential legitimation to prove that the pay gap does not exist. Here, other barriers contributing to the pay gap are positioned as *misinformation and cherry picked data*. The implication is that if others are under-paid it is because, unlike the relatives listed, they are not *hard working women*. Here the pay gap is positioned not just as bogus, but being *shaped* and *cherry picked* suggesting a malign agent.

## Sexual violence and the gender pay gap

### 6 Delegitimisation and the pay gap

Example 198 My wife my mother my sister all make more or equal to the males in there respective fields. They are not lawyers doctors, but they are hard working women. The PAY GAP IS A LIE shaped with misinformation and cherry picked data.  
(TwitterPG, PAY GAP IS A LIE, 2020-02)

The term *identity politics* describes activism by minority social groups to protect their joint welfare and interests, when they do not have power to act directly. The writer of Example 199 takes an antagonistic position against *all interest groups*. The first sentence dismisses the pay gap as *So much BS*, before going on to further delegitimise it with *so-called*, suggesting unjustified use of a term (Oxford English Dictionary [electronic resource], n.d.-1. In highlighting *interest group[s]* and *ID politics* the writer takes a position outside these categories, and by implication aligns with the dominant group. Hostility seems to be directed towards *feminism* as much as toward the pay gap itself. This may reflect the fear that a win for an interest group could result in a loss for the mainstream (Cole, 2007, pp. 4, 11, 23, 24, 153, 154).

Example 199 So much BS & so few characters to say it in. The so called PAY GAP IS A LIE. What you need to understand is that feminism is an interest group & like all interest groups they invent greivences & make up statistics. The most hatred comes from advocates of ID politics  
(TwitterPG, PAY GAP IS A LIE, 2020-10)

#### 6.4.5 Fake - TwitterPG

Within NexisPG, a WordSketch finds no occurrences of *fake pay gap*. In contrast, TwitterPG has 21 (LogDice 8.6) and all but four represent the pay gap as bogus, either accompanied by an equal pay argument, or by an insult. The remaining four challenge the idea that *the pay gap is fake*. Ten of the 21 explain that gender-based pay disparity is justified, because of: differences in: *education; experience and position; women's choices; market forces*; and a single reference to manipulation by the new world order<sup>86</sup> (NWO) (van der Tak & Harambam, 2024) to de-stabilise the status quo.

The differences in *education, experience, position, and hours worked* referenced in Example 200 would provide evidence of an equal pay failure, but are not pay gap issues. *Career segregation* is a pay gap issue, but to be *Rightful* requires a level playing field where career choices are not restricted by socially constructed discriminatory barriers.

Example 200 Gender PAY GAP IS FAKE, provide proof of same education, experience and position that has a wage gap? Rightful effect of hours worked, career segregation  
(TwitterPG, PAY GAP IS FAKE, 2019-06)

Example 201 uses different phrasing to position the pay gap as *fake*, with the agentic phrase *women are choosing* different jobs which, as in the last example, ignores any social barriers. The closing *google is free* statement is an insult, implying that the intended audience is ignorant of facts that are readily available. However, Google's search algorithm uses the searcher's browsing history to provide "tailored results based on your activity" (Google, n.d.)

---

<sup>86</sup> A long established but still popular conspiracy theory.

potentially reinforcing their existing beliefs. Often termed an echo chamber, this mechanism is thought to drive polarised and extremist thinking (Rafail & Freitas, 2019, p. 3).

Example 201 the gender PAY GAP IS FAKE, women are choosing to have different jobs  
google is free (TwitterPG, PAY GAP IS FAKE, 2021-02)

The commonsense sounding *you get paid based on what you bring in* from Example 202 calls on Rational legitimation to imply that if employees are paid less it is because they are worth less. The third sentence, *it is a business*, seems to refer to a free-market and rule-free situation that does not take account of discriminatory societal barriers. The insult in the final sentence suggests that the intended audience doesn't understand the issue because it is not paying attention.

Example 202 you get paid based on what you bring in. its very simple. it is a  
business. it is already against the law to pay based on sex or race, so i dont even  
know what you are talking about. the PAY GAP IS FAKE, and anyone who pays  
attention knows that. (TwitterPG, PAY GAP IS FAKE, 2020-02)

#### 6.4.6 False - TwitterPG

Within NexisPG a PAY GAP WordSketch finds no occurrences of *false*. In contrast, there are 21 *false* concordances within TwitterPG, and of those, eight provide a justification including: *factors of employment; jobs and experience; [having] children; working less hours; overtime; [men doing] more dangerous work*. These refer to equal pay rather than the pay gap issues.

Example 203 demonstrates the economical use of space in providing the statement, justification and an insult all in 13 words. By referring to *Research* the writer gains an element of Scientific legitimation. However, if research relies on internet searches there is a risk that it will only reinforce existing beliefs.

Example 203 PAY GAP IS FALSE. Research comparable jobs and experience... see? No  
pay gap. Morons (TwitterPG, PAY GAP IS FALSE, 2019-10)

In an unusual tweet, Example 204 claims the perceived falsehood is promoted as *just another narrative to divide us*. According to the idiom "divide and rule", the purpose of dividing people is to prevent them acting on a common cause against those ruling (Galer, 2020; Oxford English Dictionary [electronic resource], n.d.-m). In the context of this study, this may refer to a conspiracy theory.

Example 204 The gender PAY GAP IS FALSE. It doesn't exist. It's just another narrative  
to divide us. (TwitterPG, PAY GAP IS FALSE, 2020-11)

Example 205 is discussing an article written by a woman, the *pick me girl*, who is denying the pay gap. Being a woman speaking on a women's issue the *pick me girl* gains Personal Authority, potentially giving her words more power than the same text written by a man. The final sentence is sarcastic, and positions *pick me girls*<sup>87</sup> as betraying womankind.

---

<sup>87</sup> A *pick me girl* is an adult woman who says she is "different from other women" and purposefully adopts male values to ingratiate herself, seeking attention or favour. (Urban Dictionary, n.d.-c)

Example 205 Reading an article on why the GENDER PAY GAP IS FALSE !! And it’s written by a women !!! Don’t you love pick me girls !!!

(TwitterPG, PAY GAP IS FALSE, 2021-01)

**6.4.7 True - TwitterPG**

Within NexisPG *true* never collocates with PAY GAP, while TwitterPG returns 13 *true* concordances (LogDice 7.8). *True* may sound supportive and positive, but closer inspection of the concordances, shown in Table 66, revealed that of the 13, only six include the phrase the *pay gap is true*. Four state directly that it is *not true*, and three discuss a hypothetical *pay gap* using *if* or *assuming*.

For the six *pay gap is true* concordances, the wider text shows that the meaning is inverted. Two are sarcastic, two discuss the *belief* that the *pay gap is true*, and one is part of a list of left-wing stereotypes intended to insult a politician.

Use of <i>true</i>	Frequency	Percentage
Total of 13		
<i>is true</i>	6	46%
<i>not true/Isn’t true</i>	4	31%
<i>if/assuming the pay gap was true</i>	3	23%

Table 66, Manual breakdown of the adjective predicate *true* within TwitterPG.

None of the *true* tweets represent the pay gap as social wrong.

The apparently straightforward statement that the *pay gap is true after all* in Example 206 is revealed as ironic when viewed alongside the photograph that it links to, displayed as Figure 23.

Example 206 Gender PAY GAP IS TRUE after all <https://t.co/5SGufoBOYf>

(TwitterPG, PAY GAP IS TRUE, 2020-05)

The US women’s soccer team’s law suit ran for six years before it was settled in 2022, two years after this tweet was written. The tweet in the snapshot of Figure 23 refers to Judge Klausner dismissing the US soccer team’s claim for back pay. The photo of a celebrating Rapinoe is used as a contrast against the dismissal of the case, creating a heightened sense of downfall and humiliation. In 2022 the US women’s national team were awarded \$24 million in back pay and the pledge to equalise pay with the men’s team from the next round of pay agreements.

Figure 23, Tweet linked from Example 206  
Megan Rapinoe, 2019 FIFA World Cup.

<https://t.co/5SGufoBOYf>



The ever-popular discussion of politics includes a small number of tweets that can be identified as being of UK origin. As in the US, the pay gap in Example 207 is positioned as a stereotypical left-wing issue. The most likely target of this abuse is Nicola Sturgeon, then Scotland's first minister, who was attracting negative headlines at the time. The descriptors listed are comprehensive, from politics, to entertainment preferences, by way of physical attributes. This seems to convey not just disagreement with, but disapproval of every aspect of the individual. Here the *pay gap is true* carries negative prosody, and is used to delegitimise in the same way that someone might use the term flat-earth<sup>88</sup>.

---

<sup>88</sup> Flat-earth<sup>er</sup> a person who continues to believe that the planet Earth is flat despite the evidence to the contrary (Merriam-Webster [electronic resource], n.d.)

Sexual violence and the gender pay gap  
6 Delegitimisation and the pay gap

Example 207 @Chr1sMcLaughlin @murdo\_fraser @NicolaSturgeon Without looking....leftie, independent, hates Tories, hates Brixit, suddenly loves the EU, no borders, feminism, PAY GAP IS TRUE, free education, slightly overweight, not as many friends, watched braveheart<sup>89</sup> 10times. You bore the life out of me.  
Cheers (TwitterPG, PAY GAP IS TRUE, 2020-09)

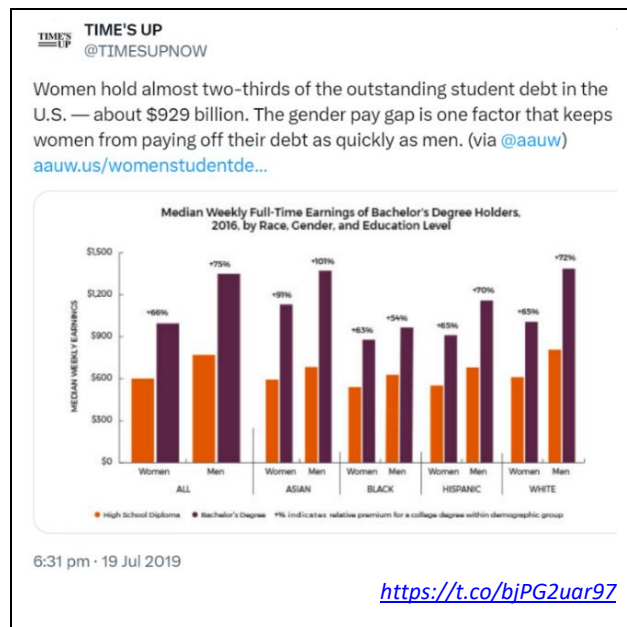
### 6.4.8 Hoax - TwitterPG

Within NexisPG, WordSketch finds no occurrences of *pay gap is a hoax*. There are however eleven *hoax* (logDice 7.7) concordances in TwitterPG, with two calling out pay gap denial. Four specific arguments arise: *more women going to university* (x 2); *non-monetary benefits*; *taking low paid jobs*; and *woman have children which disproportionately takes them out the workplace*. The remainder of the concordances are more ambiguous.

The writer of Example 208 states *it is illegal to pay women less*, confusing the pay gap with equal pay, and also assuming that the existence of legislation guarantees compliance.

They also discredit the AAUW (American Association of University Women, 2015, p. 28) graph, shown in Figure 24, as *THE STUPIDEST of the week*. The writer uses commonsense Rational legitimation that *MORE WOMEN ATTEND COLLEGE THAN MEN*. It is true that in 2019 women made up 57% of undergraduate students in the US (National Center for Education Statistics, 2021). However, 70% of US students work, and women affected by the gender pay gap earn less while studying and may require larger loans, and after gaining their qualification the pay gap may slow their repayments, extending the load period and increasing their debt (American Association of University Women, 2015, pp. 17, 18, 19, 2021, p. 4). Additionally in 2021 1.7 million female students in the US were single mothers, potentially incurring childcare costs more likely to affect women.

Figure 24, Tweet “Women hold 2/3<sup>rd</sup> of US student debt”



Example 208 This has to be THE STUPIDEST graph of the week. First of all, it is illegal to pay women less and has been for awhile, the gender PAY GAP IS A HOAX. Secondly, MORE WOMEN ATTEND COLLEGE THAN MEN. #genderpaygap #college <https://t.co/bjPG2uar97> (TwitterPG, PAY GAP IS HOAX, 2019-07)

An alternative approach is taken in Example 209, which directly addresses the US politician Alexandria Ocasio-Cortez (@AOC). *Non-monetary benefits* is an unusual justification that is not included in any reporting, and it isn't obvious how this would be evaluated. The last

<sup>89</sup> Braveheart is a movie about 13<sup>th</sup> century Scottish warrior William Wallace.

sentence advising *DON'T TAKE IT* attributes agency to a candidate, but to act would require that the candidate knows that they have been undervalued, and that they have another option. This represents the underpaid as culpable for their disadvantage, in effect victim-blaming.

Example 209 @AOC thus it's not a gap. "pay" is not just what is in your check, it's also the non monetary benefits you receive as well. When you add it all up the PAY GAP IS A HOAX. Don't want the job? DON'T TAKE IT

(TwitterPG, PAY GAP IS A HOAX, 2019-05)

Gould et al (2016, p. 4) found that in addition to wage disparity, women were also likely to receive lower health insurance, pension benefits, bonuses, and shares. A poorer benefits package could be costed and would widen the pay gap rather than closing it.

## 6.5 Measure of Acceptability

From the earlier analysis based on Table 62, the TwitterPG Acceptability concordances generally represent the pay gap as unacceptable and needing to be closed. Only a small percentage of the TwitterPG Acceptability concordances represent the pay gap as acceptable and these use: *bad* (3% of 68); *wrong* (10% of 20); *important* (11% of 18); and *problem* (26% of 50). The exception is *fair* where the majority (73% of 15) believe the pay gap is acceptable but justified.

NexisPG contains a single Acceptability adjective predicate, *bad* (LogDice 10.1) and they all represent the pay gap as unacceptable and in need of a resolution.

These TwitterPG concordances are discussed in section 6.5.1, and those of NexisPG in section 6.5.2.

### 6.5.1 *Bad* - TwitterPG

Unusually, *bad* collocates with PAY GAP in both the TwitterPG (LogDice 10.0) and NexisPG (LogDice 10.1) corpora. Within TwitterPG the strongest collocate of *bad* is its comparative form *worse* (LogDice 11.03), followed by the intensifiers *even* (LogDice 7.05) and *much* (LogDice 5.7). Unusually only two *bad* concordances dispute the existence of the pay gap, and only two concern US politics. The remainder illustrate straightforward statements of Moral Evaluation of the pay gap as unacceptable. As the pay gap is a population level statistic, differences between sub-groups are expected, and within the TwitterPG concordances 68% represent the pay gap as *worse* for different intersectional demographics and industries, and as made *worse* by the Covid19 pandemic. *Bad* concordances predominantly represent the pay gap as existing and unacceptable, while *worse* may be used to introduce another issue, but without devaluing the pay gap.

In Example 210 *women of color* are foregrounded, while *woman* and *man* are unmarked, but from the context they must be white women and men. It is a common criticism that public debate foregrounds white demographics when *women of color* and other demographics may additionally experience racial, religious, LGBTQ+, or other types of discrimination.

Example 210 Today is about the average of the 80 cents a woman makes compared to a man, but the PAY GAP IS EVEN WORSE for most women of color. #EqualPayDay  
<https://t.co/JJWHL7jDIw> (TwitterPG, PAY GAP IS BAD, 2019-04)

The statement in Example 211 that the *pay gap is bad* calls out *snoop dog* who has advocated in support of closing the pay gap on Instagram. As a US music celebrity, Snoop Dog is an influential Role Model, but here he is accused of *spread[ing] miss information* and his followers are discredited for *follow[ing] like sheep and do[ing] 0 research at all*. This is the only *bad* concordance to refer to *research*, which is more common in other themes.

Example 211 Yeah people just spread miss information like snoop dog posted a Instagram vid saying how the PAY GAP IS BAD so a lot of his fans fueling the fire so sad to see people follow like sheep and do 0 research at all  
(TwitterPG, PAY GAP IS BAD, 2019-07)

### 6.5.2 *Bad* - NexisPG

Of the 14 NexisPG *bad* concordances ten used the comparative form *worse/worst*. Those worse off include: *church*, *older women x 2*, *Cornwall*, *skilled women (India)*, *estranged wives*, *BBC x 2*, *women in Scotland*, *women in our fair city*, *finance and insurance firms*, and *women of color*.

Example 212 is a headline from The Times stating *The gender pay gap is bad enough*, suggesting that the situation has reached the boundary of tolerability (Oxford English Dictionary, n.d.-n). However, this limit has been breached by *estranged wives [who] have it even tougher*, referring to the pension gap that divorced women may also be exposed to. Generally earning less, and with fewer years of employment and paying into pensions due to child rearing, women generally have less pension provision in their own names. Here the plight of *estranged wives* is raised in a way that adds gravity to the problem of the pay gap, rather than detracting from it.

Example 212 The hidden pension cost of divorce; The gender PAY GAP IS BAD enough, but estranged wives have it even tougher  
The Times (London) (NexisPG, PAY GAP IS BAD, 2020-06)

### 6.5.3 *Wrong* - TwitterPG

Within NexisPG, WordSketch finds no examples of *wrong* as a collocate of PAY GAP. In contrast TwitterPG has 20 *wrong* collocates with a LogDice of 8.4, and the majority of these use it as a Moral Evaluation, to indicate the pay gap is unacceptable, with only two examples discrediting the pay gap.

The 12 concordances with the phrase the *pay gap is wrong* acknowledge the existence of the pay gap which validates the individuals concerned, and this is illustrated in Example 213. Here, the use of the first name, *Hannah*, creates a personal connection, and the closing *All the best to you* makes this tweet unusually supportive for this corpus. The writer uses the modal *should* to suggest that this ought to be the case (Oxford English Dictionary, n.d.-G) although it may not reflect current reality.

Example 213 Hannah. Well done to you. We're your bosses male. If so it would have been more difficult. This gender PAY GAP IS WRONG. pay should be by scale or by merit. All the best to you (TwitterPG, PAY GAP IS WRONG, 2020-10)

Example 214 calls out pay gap deniers, first by representing *Actual statisticians* as having Expert and Traditional Authority through their relationship with *government / big companies*. This is contrast with a *Random white man with no stats skill* who, nevertheless tries to invoke Scientific legitimation using words such as *statistically*. The final sentence is an equal pay phrase commonly associated with the stereotype, *Women get paid the same amount for the same roles!*.

Example 214 Actual statisticians who work the government / big companies publishing gender pay gap report: Random white man with no stats skill: the GENDER PAY GAP IS STATISTICALLY WRONG. Women get paid the same amount for the same roles! (TwitterPG, PAY GAP IS WRONG, 2019-06)

This tweet is most likely intended as humour.

#### **6.5.4 Important - TwitterPG**

Within TwitterPG an acknowledgement that the pay gap is *important* is often used to introduce another issue, such as (paraphrasing): not discriminating against anybody; discrimination in the workforce; rape gangs; FGM<sup>90</sup>; the chance gap; abusive relationships; and total equality. While the NexisPG corpus does have *important* (LogDice 8.6) collocating with PAY GAP, it falls below the frequency threshold so is not discussed here.

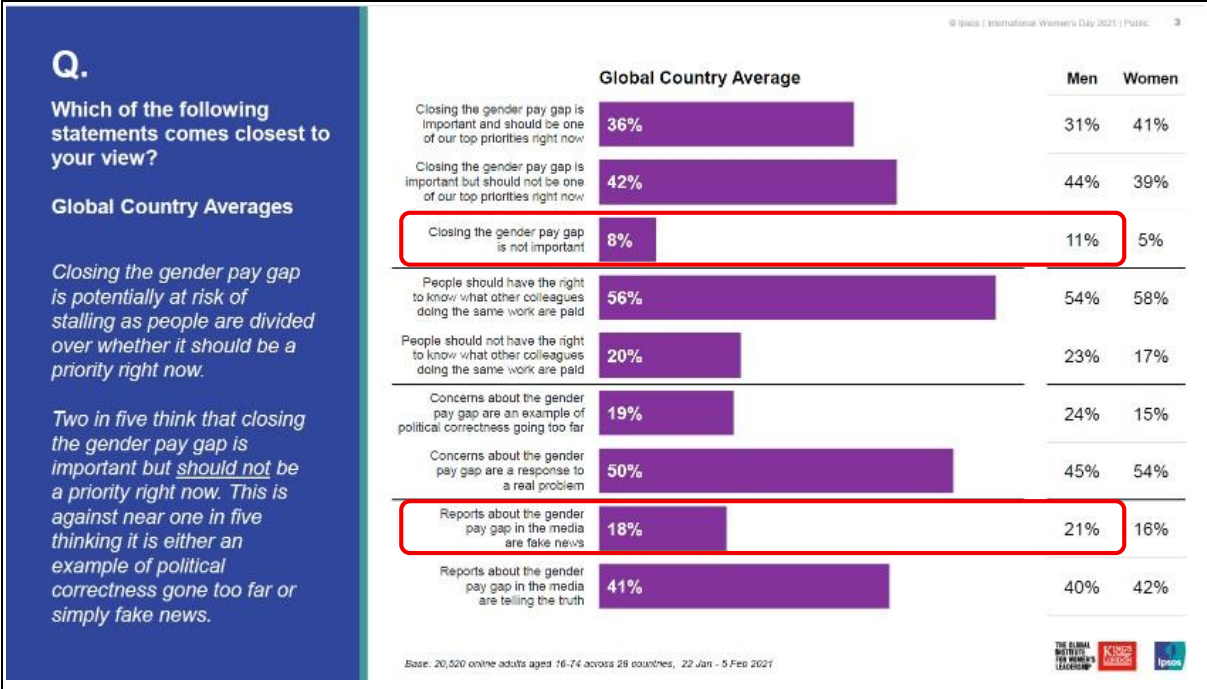
None of the 18 (LogDice 8.2) TwitterPG concordances deny the existence of the pay gap or its importance, however two quote the phrase *the pay gap is not important* from a survey question in the 2021 International Women's Day Report (The Global Institute for Women's Leadership & Kings College London, 2021, p. 3) a page from which is shown in Figure 25.

---

<sup>90</sup> FGM, female genital mutilation

Sexual violence and the gender pay gap  
 6 Delegitimisation and the pay gap

Figure 25, International Women’s Day Rreport  
 A fifth (21%) of men think media reports about gender pay gap are ‘fake news’.



The report attracted media attention, possibly because of its association with the high profile partners Ipsos Mori and King’s College London, giving the tweet in Example 215 Traditional and also Expert Authority. The tweet uses the eye-catching phrase *one in 10 men said tackling the pay gap was not important*, referring to a Canadian TV News article, shown in Figure 25, rather than the inverse but less dramatic 75% (31% favouring immediate closure, and 44% viewing it as not a current top priority) of men who thought it was important. Combined with the ironic *Oh, yeah, #HappyInternationalWomensDay*, the quote is designed as a hook to attract people to view the article.

Example 215 More than one in 10 men said tackling the PAY GAP WAS NOT IMPORTANT, according to the poll published on International Women’s Day by Ipsos MORI and the Global Institute for Women’s Leadership (GIWL) at King’s College London. Oh, yeah, #HappyInternationalWomensDay <https://t.co/8HrmmOUma0>. (TwitterPG, PAY GAP IS IMPORTANT, 2021-03)

A reply to an earlier tweet in Example 216 suggests that the writer had previously been interpreted as denying the pay gap. It is fairly representative of the *important* concordances in supporting closure of the pay gap. The second sentence states the writer’s position unambiguously *Of course the pay gap is important* before introducing another issue, general discrimination of *anybody - black, gay, religious, female, etc*. This could be a well intentioned attempt to improve visibility of intersectional demographics, but the use of *anybody* could be problematic. Broadening an anti-discrimination message from the disadvantaged group to *anybody* has a commonsense Rational legitimization appeal which superficially suggests equality. However, it also has the effect of removing from view the socially constructed barriers hindering particular intersectional demographics such as *black, gay, or religious* communities. When applied to ethnic groups this is colour-blindness, which may be unconscious, but is associated with racism (West et al., 2021, p. 1139 1140).

Example 216 I'm not saying that all, neither have I implied it. Of course the PAY GAP IS IMPORTANT, as is not discriminating against anybody - black, gay, religious, female, etc. (TwitterPG\_01to24, PAY GAP IS IMPORTANT, 2019-06)

A different view of intersectionality, where individuals may be part of multiple demographics (Government Equalities Office, 2019, p. 3), is given in Example 217. The writer points out the irony that those in the best position to pursue discrimination cases, and most frequently discussed, are upper middle-class and white, who suffer least discrimination. That the writer needs to say this, suggests an awareness that some readers are likely to be blind to the discrimination of other intersectional groups.

Example 217 Yes, addressing the PAY GAP IS IMPORTANT. But you're mostly upper middle-class and white and cling HEAVILY to those identities to prop yourself up. The only people LESS oppressed than you specifically are the men in your demographic, Barbra. (TwitterPG\_01to24, PAY GAP IS IMPORTANT, 2020-08)

### 6.5.5 Fair - TwitterPG

Of the 15 *fair* concordances only four position the pay gap as discriminatory, with ten stating that it is not only fair, but justified because: *there are good reasons; Many go for jobs in care + work less hours; you can't make money come from nothing; utter bollocks!; and the market ... determined your wage*. The strongest themes for the *fair* concordances concern women's choices and football.

Women's choices are the subject of Example 218 where the writer states that the *pay gap is fair*, and uses the commonsense Experiential legitimation *Women don't go to work on oil rigs*, and *Many go for jobs in care* to justify this. The reference to the *exact same jobs* confirms that the writer supports equal pay, but this could lead readers to conclude women make career choices unconstrained by systemic barriers such as unpaid care burdens, and biased workplace processes.

Example 218 @ChukaUmunna @CranfieldUni I am against companies that pay men more than women for the exact same jobs. Women don't go work on oil rigs, many that go thru the rigours of corporate escalation decide it's not worth it. Many go for jobs in care + work less hours. The PAY GAP IS FAIR, not patriarchal. (TwitterPG, PAY GAP IS FAIR, 2019-07)

Figure 26, Alex Morgan  
Linked from Example 219

The US national women's soccer team fought a bitter six year law suit over equal pay and conditions. Example 219 counters the frequent accusation that the women's poor pay was commensurate with the quality of their game. It responds to the linked tweet, shown as Figure 26, claiming Experiential legitimisation arising from their 2015 Women's World Cup win, football's greatest prize.



Example 219 anyone who claims that the quality of the game play is not comparable and therefore the PAY GAP IS FAIR: you haven't watched the game  
<https://t.co/9tpFUaE0eQ> (TwitterPG, PAY GAP IS FAIR, 2019-07)

### 6.5.6 Problem - TwitterPG

Of the 50 *pay gap is a problem* concordances, 13 (26%) state that it *isn't* (LogDice 5.19) or is *not* (LogDice 3.01) a problem. Additionally, twenty (40%) of the 50 attribute responsibility either for causing or resolving the pay gap to different levels: systemic (1); global (4); national (6); industry (5); company (2); and the individual (2).

Korea is rarely mentioned in these corpora, and lack of debate may reflect their paucity of equal pay legislation, which for most countries needs to be a precursor of pay gap reduction. Example 220 gains Traditional Authority by referencing the *OECD* in calling out Korea's pay gap as a *serious problem* and *unreasonable*. The writer recognises the pay gap as socially constructed and therefore solvable, if the government enacts *policies to limit discrimination*. The *wage gap* refers to equal pay which aligns with the phrase *same labor*.

Example 220 PAY GAPS ARE A SERIOUS PROBLEM, especially in Korea. The wage gap between men and women in Korea is the largest among OECD countries. It is unreasonable for a woman to be discriminated against despite the same labor. Policies to limit discrimination are needed.  
(TwitterPG\_01to24, PAY GAP IS A PROBLEM, 2019-04)

In Example 221 the pay gap is also identified as a *systemic problem*, and the writer responds to a common justification (referred to previously in section 5.4.1.1) that women fail to get fair pay because they lack *confidence* to negotiate better contracts. The writer has

Experiential legitimization in speaking for @Lighthouse which upskills disabled workforces through rehabilitation.

Example 221 GENDER PAY GAP IS A SYSTEMATIC PROBLEM and not a confidence problem with one gender. We @AT\_Lighthouse believe with right skills everyone is equally relevant & should be paid equally. Thanks @Ellevest for highlighting that. @SallieKrawcheck #StayRelevantForever <https://t.co/2l8EXpUSsF>

(TwitterPG, PAY GAP IS A PROBLEM, 2019-04)

The tweet links to Figure 27 which is a video of Sallie Krawcheck, the former head of Bank of America's Global Wealth and Investment Management division, and CEO of Ellevest, who has been called "the most powerful woman on Wall Street" (Wikipedia, n.d.-b). The two images in Figure 27 show Ms Krawcheck, refuting the claim that the gender pay gap is a women's confidence issue. She explains that women's and men's confidence is perceived differently, with assertive confident women viewed as aggressive and unpleasant. Cooke highlights a cultural asymmetry in which women, increasingly since the 1970s, have identified with stereotypical male traits, but men's identification with female traits was minimal (2024, p. 355).

Figure 27, Tweet "No, no, no. Of course its not. It's a systemic issue."



## 6.6 Contempt

There are four words collocating with PAY GAP in Table 62 that suggest contempt: *bullshit* (LogDice 9.2), *nonsense* (LogDice 8.3), *ridiculous* (LogDice 8.2), and *joke* (LogDice 7.4). This

category includes examples that show contempt for both an acknowledged and wrongful pay gap, and also for a bogus pay gap.

### 6.6.1 *Bullshit* - TwitterPG

NexisPG returned no PAY GAP concordances referring to *bullshit*, in contrast to 31 that state that the *pay gap is bullshit* within TwitterPG, and 30 of those deny its existence. The most frequent sub-theme, with five concordances, is discrimination against men, while four other concordances remind readers that equal pay legislation exists.

It is implied that the author of Example 222 is a woman by the use of *we're*, giving the tweet Personal Authority. The writer states that the *pay gap is bullshit*, distancing herself from other women. Having just said she is not like *all women* she re-establishes her femininity with the humorous *it makes me moody*, referring to the moody woman stereotype.

The writer also establishes her political credentials with *conservative and believe the pay gap is bullshit*, two attributes frequently co-occurring in this corpus. Flattering her intended audience with the epithet *chief*, and positioning herself as aligned with their values, the writer is reminiscent of the *pick me girl* trope discussed with Example 205.

Example 222 Not all women. I'm a conservative and believe the PAY GAP IS BULLSHIT and there's a war on men. Stop assuming we're the same, chief - it makes me moody. (TwitterPG, PAY GAP IS A BULLSHIT, 2019-10)

Criado-Perez (BBC, 2014) suggested that women may disaffiliate from feminist issues because they have absorbed the hegemonic culture (section 2.3.2).

Several of the *bullshit* concordances are brief, like Example 223 which refers readers to *Jordan Peterson*, invoking his Expert and Scientific legitimation. Peterson is a popular but controversial psychologist of the right and a media figure, who views men and women as having natural differences that result in gendered and appropriate life choices (Peterson, 2018). However, his arguments generally omit references to gendered cultural constraints.

Example 223 See Jordan Peterson. Pay gap is BULLSHIT (TwitterPG, PAY GAP IS A BULLSHIT, 2020-06)

The name *Peterson*, occurs 145 times in the TwitterPG corpus, making him the most quoted source. He is referenced exclusively in denials or justification of the pay gap. In contrast the name appears only 13 times in NexisPG corpus.

### 6.6.2 *Nonsense* - TwitterPG

Sixteen of the 17 adjective predicate *nonsense* concordances deny the existence of the PAY GAP, but these are quite varied, with four concerning equal pay, two discussing politics, and the remainder unrelated or only loosely concerning the pay gap.

Despite the initial phrase featuring the pay gap, the topic of Example 224 is equal pay. This is a good example of the commonly mis-aligned debate where the writer focuses on equal pay issues such as, *same job same pay*, while the imagined interlocuters, *feminists*, are discussing the wider *pay gap*. Feminists are delegitimised here with reverse Scientific legitimation, for failing to *do math and basic logic*.

Example 224 This PAY GAP IS NONSENSE, well at least in Australia is illegal to pay different rates. Same job same pay. Don't any feminists do math and basic logic? (TwitterPG, PAY GAP IS A NONSENSE, 2019-08)

With its political theme, Example 225 clearly aligns denial of the pay gap with *the right*. Here, *plenty of people* suggests Authority through Conformity to *anyone on the right*, while the phrase *they can tell you* carries the presumption that all people *on the right* are well informed on the pay gap and of the same mind. In contrast, the intended audience is subject to Experiential delegitimation with *You are just in a bubble*, implying isolation from reality (Oxford English Dictionary [electronic resource], n.d.-i).

Example 225 Also, there are plenty of people who disagree with this political ideology. Just look at anyone on the right and they can tell you all about how equality of outcome and the gender PAY GAP ARE NONSENSE. You are just in a bubble. (TwitterPG, PAY GAP IS A NONSENSE, 2020-03)

### 6.6.3 *Ridiculous* - TwitterPG

Among the 17 concordances including the adjective predicate *ridiculous*, there are eight stating that there is a pay gap and it is unacceptable, while four either deny its existence or believe it is justified. The remaining five concordances are ambiguous. Of the 17, six are in the context of sport.

The BBC's scandal was an equal pay issue, although it may feed into a pay gap issues such as recruitment, promotion and career segregation were not part of the discussion. Widely reported as a pay gap, as it is in Example 226, the BBC is positioned as responsible for a *ridiculous* pay gap (Oxford English Dictionary [electronic resource], n.d.-C). Their culpability is emphasised by creating a contrast with their reputation for political even-handedness. The *tripping over themselves* suggests they are "very eager" (Cambridge Dictionary [electronic resource], n.d.-a) to be seen as un-biased, with the bracketing of (*claiming*) highlighting the possibility of failing to deliver on that claim.

Example 226 You see the BBC tripping over themselves (claiming) to show balance in the news, but their gender PAY GAP IS RIDICULOUS. (TwitterPG, PAY GAP IS RIDICULOUS, 2019-10)

Multiple Moral Evaluations are given in Example 227 starting with *All this rubbish* to describe the pay gap. The writer then represents the standard of women's football as not *coming within a million miles of the men's game*, despite their two World Cup victories. The final sentence seeks to shut down the debate with *until then*, which sets up a further but vague threshold. The six year legal action taken by the US national women's team to obtain fair pay and conditions is trivialised here as *moaning about pay*. This tweet was written a month before the US won the Women's World Cup for the fourth time.

Example 227 All this rubbish about the PAY GAP IS RIDICULOUS. When the women's game comes within a million miles of the men's game they can start moaning about pay. Until then, just shut up and play. (TwitterPG, PAY GAP IS RIDICULOUS, 2019-06)

Alkhamash (2021) analysed tweets discussing the US women's football team using sentiment analysis. In contrast to the present study, she reported no tweets denying or

justifying the pay gap within either those of positive or negative sentiment. However, all of her data was harvested in the month following their win, which may reflect a period of greater period of positivity and celebration.

#### 6.6.4 Joke - TwitterPG

Within TwitterPG the *Pay gap is a joke* (LogDice 7.4) appears in nine concordances, four of these say that there is a pay gap and it's wrong, while two state that it is bogus. The remaining three are so brief as to be ambiguous.

The writer of Example 228 establishes themselves with Personal Authority and Experiential legitimation by aligning with second wave feminism; *I have respect for the previous feminist*. From this moral high ground, they criticise *modern* ones as a *snare*, suggesting that they may be creating a trap. Although the writer acknowledges that the pay gap operates at a *statistical* level, the criticism that *they don't account for hours worked, positions or Jobs*, tries to claim the commonsense Theoretical Rationalisation legitimation type. However, this is another example of equal pay arguments being used in the pay gap debate. The writer's reference to modern feminists being *so crabby* suggests Cole's fourth characteristic way of undermining victims (2007, p. 27) discussed in section 2.5.4, where victims speaking out are accused of aggressiveness.

Example 228 I was a nurse before 3<sup>rd</sup> wave feminism hit. I have respect for the previous feminist but the modern ones are so crabby they are a snare to even be around..... the statistics for the PAY GAP IS A JOKE. They don't account for hours worked, positions or Jobs. Bill gates  
(TwitterPG\_01to24, PAY GAP IS A JOKE, 2020-05)

The US national women's soccer team make another appearance in two of the nine, which both refer to their legal suit. In Example 229 the complex case involving historic back pay, bonuses, prize money and working conditions is misrepresented as a choice. In *suddenly want a different deal* the team is represented as being childish and poor losers because they chose *larger salary and smaller bonuses*.

Example 229 @mPinoe and her equal pay lawsuit is a joke. The woman chose a larger salary and smaller bonuses and then when they won the world cup they suddenly want a different deal? The men are actually paid less per game but negotiated larger bonuses. The PAY GAP IS A JOKE  
(TwitterPG, PAY GAP IS A JOKE, 2020-05)

At the time of this tweet the court case had just been dismissed because the women had earned more than the men, but this ignored the fact that the women had won the World Cup while the men were knocked out early on. Had the men won, they would have been paid vastly more. In 2022 the women's team won a settlement delivering \$24 million in back pay.

## 6.7 Conclusion

The analysis in this chapter used the full 24 month corpora to address the third research question, which analysed truthfulness and legitimacy within the pay gap corpora, and how it differs from the sexual violence corpora.

The answer to RQ3 is that truthfulness and legitimation are attributed to the phenomena of the pay gap. Individuals in pay gap concordances are delegitimised indirectly for believing and advocating for its closure. In contrast, the existence of sexual violence as a phenomenon is never disputed, and delegitimisation targets the victimised individuals. However, it is more complicated than that, as will be explained in the remainder of this section.

The news articles of the NexisPG corpus, written by people known to the news publisher and therefore more accountable, have fewer different delegitimising modifiers that occur less frequently. These provide more measured and traditional discourse compared with Twitter examples. Analysis of TwitterPG revealed high levels of delegitimation clustering into three categories: Veracity, Acceptability, and Contempt. Veracity concordances contain representations of truth or untruth, with the largest number of such collocates and highest LogDice scores found in the TwitterPG corpus. Even apparently positive terms, such as *true* and *real*, may be inverted to deny the pay gap. The Acceptability concordances are more likely to represent the pay gap as a problem that needs to be solved, although *fair* was surprisingly often associated with justification of the pay gap as fair. The category showing most acknowledgement of the pay gap is Contempt, where *ridiculous* and *joke* account for approximately half of concordances positioning the pay gap as discrimination that should be eliminated.

Compared with the sexual violence corpora which each have a maximum of one delegitimising modifier with a LogDice of at most 10.0, TwitterPG has three: *myth* (LogDice 12.9), *real* (LogDice 12.4), and *bad* (LogDice 10.0). Furthermore, TwitterPG has a wider range of delegitimising modifiers (15) than any of the sexual violence corpora. Taken together this suggests that TwitterPG is more delegitimising than any of the sexual violence corpora, and more than NexisPG.

Where pay gaps are acknowledged, they are most often positioned as equal pay situations that result from the individual's choices, which they should resolve themselves. The personal solutions suggested include: *choose a different job; get more skills; develop more experience; and negotiate more effectively*. When framed as making choices, women are credited with the agency to overcome past and present barriers, however unsurmountable. This holds the individual accountable and removes from view their barriers to success. Represented as responsible for their own success or failure, those unable to succeed may choose to remain silent rather than be victim-blamed. The same tweets often reflect a widespread assumption that where equal pay legislation exists it is automatically adhered to, however cases such as those involving Birmingham Council (2012-2024) (Headley & Rack, 2024) and Asda (2014-2025) (Butler, 2025) provide evidence to the contrary. The prioritising of individual responsibility for overcoming difficulties and its political roots are discussed in more detail in Chapter 7. The pay gap is a statistical phenomenon that operates at population level. It is complex and includes ideological issues such as career segregation, glass ceilings, and distribution of unpaid caring work. These issues hide in plain sight, as taken for granted

aspects of life manifesting as socially constructed expectations and common sense. As an ideological issue the pay gap requires a solution at the level of society, and cannot be eradicated by the actions of the individual. However, changes at societal level entail significant adjustments, and for some people a shift in world view, thinking, and behaviour. Such changes make the pay gap potentially threatening, especially to those who fear they may lose out. With so much at stake, clear definitions of equal pay and the pay gap are crucial, but absent.

Those disadvantaged by the pay gap, may slip out of sight in a variety of ways. Some don't realise they are being short changed, or do realise but have no hope of being listened to, taken seriously, or obtaining justice. Others fear retaliation if they speak out. Within the texts some that are visible are represented as associated with a non-existent pay gap, so cannot be disadvantaged and thus go missing, and these are discussed in more depth in section 7.5.

## 7 The labelled and the missing

### 7.1 Introduction

The research question driving this chapter, repeated below for convenience, draws on evidence from the preceding analysis chapters.

RQ4 How does delegitimation manifest for VICTIMS and SURVIVORS? This question is addressed in two parts.

RQ4A, Is the SURVIVOR label more beneficial than the VICTIM label for individuals, and what are the implications for society? (sections 7.3 & 7.4)

RQ4B Who are the missing? (sections 7.5, & 7.6)

This chapter comprises two parts, the first part addresses RQ4A, and the second part RQ4B. To answer RQ4A this chapter draws on evidence from all of the preceding analysis chapters based on the sexual violence datasets, to evaluate more deeply the VICTIM and SURVIVOR labels, and how they work to the advantage or disadvantage of individuals targeted in sexual violence and society more widely. Particularly, this chapter explores the possibility that this may not be a simple dichotomy. While later evidence remains in line with the RQ1 finding that VICTIM is more strongly associated than SURVIVOR with negative representations, discussed in section 7.3, it also suggests that SURVIVOR may be associated with less obvious costs, covered in section 7.4.

The second half of the chapter considers RQ4B focusing, in sections 7.5 and 7.6 respectively, on those we might have expected to find in the discourse but are missing from pay gap and sexual violence corpora. Locating the missing created a daunting challenge that Duguid and Partington aptly describe as “both a paradoxical and hopelessly wide-ranging endeavour.” (Marchi & Taylor, 2018, p. 39). However, using a process of elimination, it is possible to say something about the likely characteristics of the missing, and who they are not. Missing individuals in this study who are subject to either sexual violence or the gender pay gap include:

- Van Leeuwen’s Suppressed and Backgrounded;
- Those who don’t know they have been victimised;
- Those who know they have been victimised but dare not speak out;
- Those subject to victimisation that is normalised and socially tolerated;
- Those who do speak out but are represented as liars, trouble-makers, self-serving, or otherwise bogus.

This chapter concludes by discussing the implications that power and the political environment have on the missing and silence.

The answer to RQ4 is discussed in section 7.7.

### 7.2 Shifting definitions of VICTIM and SURVIVOR

It has become apparent that within these corpora there is a shift in meaning of both VICTIM and SURVIVOR away from traditional dictionary definitions. Traditionally the terms VICTIM and SURVIVOR were not generally considered to be related, however the sexual violence corpora of this study show the two terms used as both synonyms and antonyms. The traditional

definition of VICTIM, the term used by the legal system, builds in the understanding that the individual is disadvantaged by a harm that is current and on-going (Oxford English Dictionary [electronic resource], n.d.-gg). In contrast, SURVIVOR was defined as “one who outlives others” after a life-threatening experience (Oxford English Dictionary [electronic resource], n.d.-J) distancing the crime and its harm into the past. However, the OED updated SURVIVOR entry includes the additional sense of:

*“A person who has experienced a traumatic event or past abuse, esp. of a sexual or psychological nature.”* (n.d.-J).

This removes the life-threatening element, and broadens the scope. VICTIM too has recently acquired additional meaning, with the following text:

*“depreciatively regarded as helpless, passive, or failing to take responsibility”.* (n.d.-gg)

Its earlier VICTIM definition carried no sense of personal responsibility, nor any less agency than SURVIVOR. The Collins Dictionary (n.d.-b) goes further, now offering SURVIVOR as an antonym of VICTIM, explicitly setting them as opposites. These changes are reflected in the analyses in earlier chapters of this study where VICTIM, particularly within Twitter, is represented as more negative than SURVIVOR, and now capable of operating as an insult. Interestingly, bearing in mind the largest number of Twitter daily users are in the US (Business of Apps, 2020), Merriam-Webster does not include any such changes in meaning for VICTIM or SURVIVOR, and gives neither word as a synonym or antonym of the other. To see VICTIM and SURVIVOR as equivalent and a simple matter of choice risks over simplifying the situation. Although organisations may choose to replace the word VICTIM with SURVIVOR based on the greater agency it carries, there are other implications. The following sections will explore how the representations of VICTIM and SURVIVOR in the present study may work for or against the individual so labelled, in sections 7.3 and 7.4.

### **7.3 Pros and cons of VICTIM**

When someone who is victimised seeks a solution from the police or justice system, at some point they will be regarded as a victim, even if they do not recognise this in themselves. Fohring puts this bluntly:

*“Is it possible to experience a crime and yet not be a victim? From an objective, technical perspective, the answer is simple: no.”*  
(Fohring, 2018, p. 156)

Any legal action raised must be described in terms of VICTIM and PERPETRATOR. Claims to resilience, strength and recovery, moves the accuser further from Christie’s Ideal Victim (1986). The perpetrator’s defense team might ask why, if no lasting harm is being claimed, a request for compensation or punishment should be upheld.

The sense of weakness that is built into VICTIM is the inverse of power, and acknowledging personal powerlessness is damaging to the individual’s sense of self (Fohring, 2018, p. 157). However, weakness is associated with vulnerability, and society usually, and unconsciously, refers to the Ideal Victim trope to evaluate those deserving of its sympathy. As part of her

research into individuals' interactions with social support services Loseke (2003, p. 134) found that the identity of VICTIM needs to be associated with a denial of responsibility to remain eligible for help and sympathy. Where individualistic cultures prioritise personal responsibility and self-sufficiency, those failing to protect themselves are viewed as personally responsible, and their victimisation as a personal failure (Loseke, 2003, p. 145).

### 7.3.1 Stigmatisation of VICTIM

The TwitterMT corpus has the highest number (29) of explicit statements of aversion to the label VICTIM, discussed previously in section 4.6.1. These 29 concordances all include the text *not a victim*, and eight of these are first person statements repositioning the writer as a SURVIVOR, as illustrated in Example 230.

Example 230 On this day every year, I am reminded that I am a survivor and NOT A VICTIM #MeToo. (TwitterMT, NOT A VICTIM 2020-02)

*On this day every year* indicates a timescale long enough to be measured in years, and the potential for eventual healing. Having declared *I am a survivor*, it seems unnecessary to continue, but the writer emphasises, *and not a victim*, suggesting something significant about distancing themselves from the undesirable VICTIM status. A lab-based moral language analysis by Lewis et al (2019, p. 8 (in pdf reader)) found that lay-people formed negative perceptions of those claiming discrimination-based victimisation, whether or not the discrimination had occurred. More strangely, they found that calling it out on behalf of another resulted in the advocate also receiving negative evaluations. Lewis speculated that discrimination acts on groups and therefore violates Ideal Victim criteria, which needs the victimisation to be individualised (Lewis et al., 2019, p. 8 (in pdf reader)).

Of the women respondents in the Advance Pro Bono study (2021, p. 22) who had experienced sexual assault, 47% said that the stigma of being a VICTIM had deterred them from reporting to the police. This is in line with Fohring's (2018, p. 161) observation that the stigma around VICTIM may result in those who need help, denying any harm. She found that the victimised may actively conceal their difficulties, to defend themselves against secondary victimisation and stigmatisation. Fohring sympathised, agreeing that there are "any number of reasons for wanting to avoid being labelled as just that, a victim" (2018, p. 162).

#### 7.3.1.1 Just World Theory

It is curious that an individual who has been subject to a crime, and may have suffered harm, would receive the unsympathetic responses seen in some tweets. An explanation may come from seminal work on Just World Theory (JWT) provided by Lerner (1980, p. 10). He explains that while victim-blaming suggests an unflatteringly selfish or callous response to another's misfortune, it is a normal and necessary self-defense mechanism. JWT (section 2.4.4) allows most of us to go about our daily lives free from the fear of everyday risks. Individuals adapt their perception of events to protect this belief, and when it is violated they may experience anger, indignance or outrage. Emotional responses of this type are evident in some of the examples seen in this study.

The consequence of accepting events as random, and the individual as not responsible for their own victimisation, is un-nerving because "if it can happen to anyone, it could also be oneself." (Fohring, 2015, p. 48). JWT operates subconsciously, prompting either a reaction to

restore justice by punishing a wrong-doer, or rescuing a victim. Alternatively individuals may refer to social scripts and cost-benefit analyses, leading to evaluations of blame (Lerner, 2015, p. 206). Thus, a JWT belief positions individuals as responsible for their own victimisation for either doing, or failing to do, something to cause it. The benefit of a JWT belief is that it creates the possibility that if we do the right things, then bad events will not befall us. Putting aside the faulty logic, the popular rape myth that says “she was raped because she was wearing the wrong skirt”, is then reframed as “if I wear the right kind of skirt I won’t be raped”. However, this type of rationalisation causes individuals to over-estimate the likelihood of good things happening to them, and underestimate the risk of bad outcomes (Fohring, 2015, p. 46). JWT may operate in all areas of life, and the strength of belief increases with social standing (Westfall et al., 2019, p. 537). Generally more privileged individuals have greater belief that the world is fair in terms of success in life, and economic rewards (Westfall et al., 2019, pp. 538, 543). This theory supports a right-leaning and meritocratic world view that success in any aspect of life is earned, and those failing to achieve success or remain safe, are responsible for their own outcomes. The demographics least likely to endorse the JWT are those who have experienced greater injustice, and who are older (Westfall et al., 2019, p. 545).

### **7.3.1.2 Authority**

Chapter 4 provided evidence that VICTIM correlates more strongly than SURVIVOR with higher frequencies of damaging judgements and evaluation. This is a problem because those requiring help, like Loseke’s (1997, pp. 426, 438) recipients of charity and US state welfare, are thus exposed to moral judgements. This is amplified to a sense of moral outrage when those asking for assistance are viewed as undeserving. Loseke suggests that restricting help to only those who conform with preferred cultural values maintains and polices the status quo. When people do not know how to view a dispute because there is little empirical evidence, they may become more focused on the mental competence and moral authority of the speaker. In this situation delegitimation may undermine a victim’s claim. Such judgments are reinforced by cultural “interlocking systems of oppression” (Anderson, 2017, p. 210). Thus when those doing the judging cannot reconcile a narrative of victimisation with their cultural beliefs, they may take that as further evidence that the speaker is mistaken, so having no cause for complaint (Anderson, 2017, p. 210).

### **7.3.2 Access to support**

By definition, all individuals victimised by sexual violence crimes start as victims, so the criminal justice system requires them to be called a VICTIM. Survivorship on the other hand, distances the crime into the past, “the underlying premise being that survivors already overcame their injuries” (Cole, 2021, p. 36). For many, awareness of their own agency and strength does not reconcile with the tarnished image of VICTIM (Mahoney, 1995, p. 62). However, Mahoney also found that both authentic and pretended agency can lead to individuals downplaying their victimisation. This can lead to denial of a crime, failure to obtain support, delayed healing, under-reporting, and failure to deal with the accused.

In taking up a position of strength and agency, a victim/survivor may be reluctant to risk weakening that image by seeking support, either consciously or unconsciously. This may be exacerbated by a culture that normalises sexual violence (Nielsen et al., 2024, p. 9) clouding individuals’ understanding of the harm done to themselves and deterring them from speaking out.

In-line with the more positive associations around SURVIVOR, discussed in Chapter 4, the lemma SUPPORT collocates more strongly with *survivor* than with *victim*, illustrated in Table 67. While the differences are not great, they are consistent across all corpora. Additionally, this advantage is more pronounced within the Nexis corpora.

SUPPORT		Freq	LogDice
TwitterMT	<i>Victim</i>	60	9.0
	<i>Survivor</i>	49	9.3
TwitterSH	<i>Victim</i>	26	8.4
	<i>Survivor</i>	17	8.9
NexisMT	<i>Victim</i>	30	8.5
	<i>Survivor</i>	36	9.1
NexisSH	<i>Victim</i>	58	8.8
	<i>Survivor</i>	40	9.4

Table 67, Lemma SUPPORT and its collocation with *victim* and *survivor*

### 7.3.3 Distortion of veracity

The present study is situated in a place and time where truth itself was a hot topic. The concordances of the *truth* debate are largely concerned with US issues, particularly surrounding the 2016 Trump presidential campaign and presidency, that were epitomised in 2017 by the phrase *alternative facts*<sup>91</sup> (Swain, 2017). Anderson describes this culture as one of “pervasive misinformation, ignorance, and ideological barriers that prevent the truth of social inequality from being widely understood and accepted” (2021, p. 2). He sees an epistemological struggle involving what constitutes truth within society (2021, pp. 4, 5). In the often over-heated atmosphere of Twitter, this creates a new uncertainty and individuals may be reluctant to speak out, fearing how their claim might be evaluated against this new type of truth where “alternative” and contradictory facts can co-exist.

The conventional understanding of truth is strongly associated with the ability to communicate it, and the verbs related to communication are highlighted in blue in Table 68. *Exposing* is included as it could have been replaced by *communicating* in all 8 concordances as illustrated by Example 231.

Example 231 Thanks to #hijaako and all prior movements w the same advocacy that women are now coming forward to the light of freedom to expose their TRUTHS and stand for their rights. The world is shifting. #MeToo  
(TwitterMT, TRUTH, 2020-06)

Particularly noticeable is the variety of terms in the Twitter corpora that are associated with conveying information, and the strength of their LogDice scores.

The initial intention of the #MeToo movement was to enable those who had experienced sexual violence to speak out by saying “Me too”. However, the majority of #MeToo tweets

<sup>91</sup> On the 23/01/2017 on NBC a senior White House aide, Kellyanne Conway, described challenges to President Trump’s false claim that crowds at his inauguration as “the largest ever” as “alternative facts”.

are by third parties, either supporting or decrying claims. While truth can be spoken, it also needs to be received as truth, and where an accuser is delegitimised, discredited, or disbelieved their purpose may be thwarted.

An association between truth and speaking out is also seen to a lesser extent in the Nexis sexual violence corpora where there are fewer speech verbs with lower LogDice scores. In Table 68 the enTenTen21 is used as a comparison, and its top ten collocates are dominated by speaking out terms, suggesting that there is a general association between communication and truth.

	TwitterMT (Frequency, LogDice)	TwitterSH (Frequency, LogDice)	NexisMT (Frequency, LogDice)	NexisSH (Frequency, LogDice)	enTenTen21 (Frequency, LogDice)
1	telling (89, 10.85)	telling (40, 10.38)	telling (38, 10.37)	telling (15, 9.7)	telling (96,496, 8.65)
2	speaking (44, 9.73)	know (20, 10.28)	simple (9, 8.72)	defence (8, 9.15)	tell 143,897, (8.34)
3	speak (41, 9.2)	#UNSG* (20, 10.12)	speak (18, 8.69)	absolute 6, (9.10)	told (94,692, 7.56)
4	tell (34, 9.01)	more (21, 9.71)	tell (16, 8.66)	reconcile (5, 9.03)	speak (44,976, 7.29)
5	power (26, 8.34)	speaking (13, 8.85)	defence (8, 8.4)	whole (9, 9.02)	God's (32,041, 7.13)
6	told (18, 8.28)	speak (15, 8.49)	pleaded (6, 8.25)	speak (11, 8.51)	absolute (28,537, 7.09)
7	speaks (10, 8.05)	tell (13, 8.12)	uncomfortable (8, 8.23)	inconvenient (3, 8.33)	reconciliation (21,902, 7.04)
8	spoke (10, 7.90)	prevail (4, 8.11)	further (8, 7.93)	disconcerting (3, 8.33)	about (31,819, 6.84)
9	comes (12, 7.80)	dying (4, 8.04)	sad (5, 7.82)	hopefully (3, 8.3)	justice (28,205, 6.80)
10	exposing (8, 7.77)	Speak (4, 8.49)	gospel (4, 7.82)	pursuit (3, 8.25)	speaking (24,998, 6.68)

Table 68, Collocations of *TRUTH* in the sexual violence corpora, and enTenTen21.

Blue cells highlight verbs related to communication.

\* UN Secretary General

### 7.3.4 Personal responsibility as victim-blaming

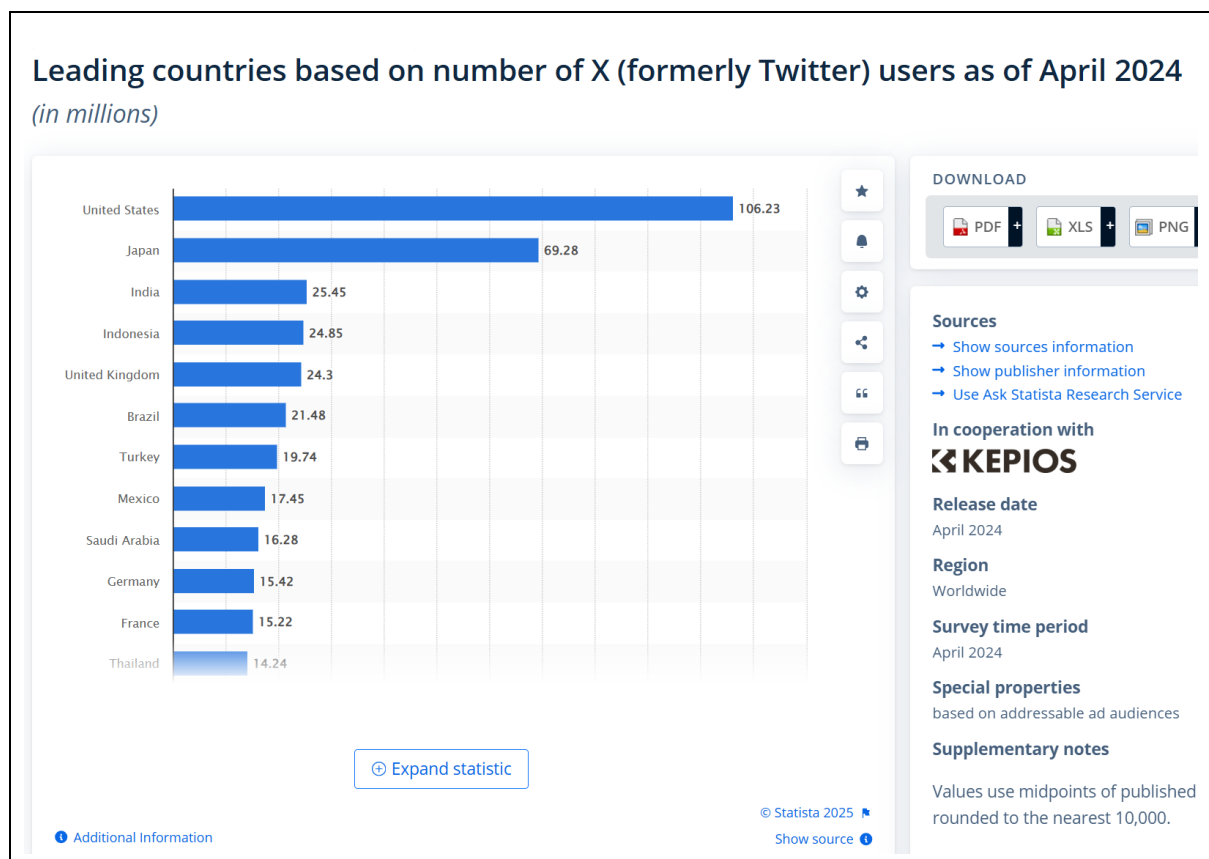
In Chapter 3 Table 20 highlighted the strong collocations of *blame* and *shame* with *VICTIM*, in contrast with its almost total absence as a collocate of *SURVIVOR* (excepting a single reference to *shame* in the NexisSH\_1Yr). *Blaming* and *shaming* will now be explored, beginning with a little context.

Following the 1960s and 70s civil rights and feminism movements in the US, there was a swing away from collective activism and towards the political right and increasing emphasis on individualism and personal responsibility (Cole, 2007, pp. 11, 12, 19, 23, 169). Loseke (2003, p. 64) defines *individualism* as “the responsibility to be economically self-reliant and

care for the self”, and views it as a primary cultural theme in the US. This shift from collective to individual recasts the weakness of a *victim* as a personal choice, and victimhood becomes a personal failure. In a later study Favaro and Gill (2019, p. 155) go further, identifying neoliberal principles as constructing individuals as responsible for their own outcomes, irrespective of cultural and systemic barriers that may operate through the “trend towards self-help culture”. Their study found that the self-help, positivity, confidence, and resilience promoted by the US therapy sector and women’s magazines often sit alongside the advertisements for the very self-improvement products required (Favaro & Gill, 2019, p. 158). They proposed that this focus on the responsibility of the individual to re-make and improve themselves could contribute to a “silencing of critique of structural inequality” (Favaro & Gill, 2019, p. 161). For example, an individual may feel their preferred self-reliant identity was jeopardised by asking for help, resulting in a reluctance to do so.

The unkindness seen in some of the *VICTIM* concordances of these corpora may be partly understood in the cultural context of Twitter’s dominant user-base, the US. Statista (Statistica, 2024) reports, in Figure 28, that the majority of Twitter users are in the US where support for individualism and personal responsibility is more common than in many other countries (Delker et al., 2020, p. 2 (in pdf reader); Loseke, 2003, p. 64).

Figure 28, Users in the United States predominate Twitter



(Statistica, 2024)

Within an individualistic culture people are encouraged to be “economically self-reliant and care for self”, which is more in line with the “strong and resilient” *SURVIVOR* than a *weak, vulnerable* and *powerless* *VICTIM* (Fohring, 2018, pp. 157, 158). However, the ability to

recover may also reflect the resources and support the victim has available (Fohring, 2018, p. 160), consequently those already disadvantaged may suffer most harm (Dillenburg, 2008, p. 13). Taken together this victim-blames and punishes those least equipped to recover (Fohring, 2018, p. 161). At the same time, organisations supporting those subject to sexual violence are increasing their use of SURVIVOR (Messamore & Paxton, 2021, p. 16). Often this is positioned as a choice free of constraints, and becoming a SURVIVOR merely an act of will power (Fohring, 2018, p. 160) (Cole, 2021, p. 36).

Even where it is possible, the shunning of VICTIM is evidence of the problematic and stigmatised nature of the label. Journalist and victim Campoamor (Campoamor, 2018, pp. 2, 3, 4) describes VICTIM as having become “bastardized” and used to “diminish, discredit, and disparage”, while SURVIVORS are “lauded as superhuman beings”. The price of adopting the SURVIVOR label is an associated obligation to provide a happy-ending story so that society can feel more comfortable about a “systemic problem”.

The legal system is based on the concept of VICTIM, which unlike SURVIVOR, is strongly associated with a perpetrator. The term SURVIVOR distances the crime into the past, reducing or removing the harm, and potentially any justice or support along with it. Some victims/survivors do not report because they feel partly to blame, but “Importantly, defining oneself as a VICTIM shifts blame to others, thus sometimes enabling a person to take action she might otherwise hesitate to.” (Dunn, 2008, p. 4 (in pdf reader)). Blamelessness is important to Christie’s (1986) Ideal Victim concept, discussed in section 2.5.1, where innocence is mandated.

The tweets in Example 232 and Example 233 both discuss the Harvey Weinstein case, but from opposing view points, illustrating the backlash that can follow naming a perpetrator. In Example 232 the *actresses* are positioned as having had no control over their abuse, but have become empowered by #MeToo to speak out. In Example 233 they are victim-blamed for their predicament, as they agentically *sold their bodies*, suggesting a quid pro quo transaction which they now renege on, because their careers are faltering.

Example 232 Fear of retribution, revenge. The actresses who said ‘no’ or spoke out against Harvey Weinstein were blacklisted. At the end of the day, CC needed to work. She made excuses & put up with it because she felt she didn’t have a choice. #METOO made it possible to speak out. Greater agency of survivor, one who has survived, not the victim of something. (TwitterMT #METOO, 2021-02)

Example 233 Who I dont feel sorry for are the adult actresses of #METOO who were adults that sold thier bodies for acting gigs got paid and are now crying rape because they stopped forty and thier roles bottomed out like thier asses and they realize they are nothing more than prostitutes. (TwitterMT, #METOO, 2019-04)

Some individuals may self-blame a particular behaviour for their victimisation, unconsciously adopting a JWT belief to protect their view of the world as a generally safe place (Fohring, 2018, p. 161). In merely changing their behaviour in the future, the individual may believe they have reduced the chances of being victimised again.

### 7.3.5 Biased use of *alleged*

In Table 13 *alleged* was seen to collocate much more strongly with VICTIM than SURVIVOR. *Alleged* concordances discuss themes of legal proceedings and US politics, but additionally may introduce of an element of cynicism or doubt. Usually, the doubt introduced by *alleged* is implied, but Example 234 is a third party report which explicitly casts doubt.

Example 234 Like RK’s attorney said, “perhaps... these ALLEGED VICTIMS are not victims at all, but only women who have been told and instructed... years later, that the claimed relationship they freely and voluntarily engaged in, should now, in the #metoo era, be classified as ‘abusive’

(TwitterMT\_01to24, ALLEGED VICTIM, 2020-03)

Here the victims of RK are represented as weak and passive in being *told and instructed*, but at the same time manipulative because their relationships *were freely and voluntarily engaged in*. In addition to its everyday use to cast doubt and criticise, a legal defense team can also use *alleged* to weaken and counter-accuse a claimant. While the collocation of VICTIM and *alleged* can mean a neutral “unproven”, here it becomes more damaging. Neither use is associated with SURVIVOR in any of the corpora.

The phrase *alleged rape* increased in the media from about 2010 as a move towards more cautious reporting after a number of high profile defendants sued their accusers for defamation (Tranchese, 2024, p. 164). Tranchese highlighted an imbalance in news media phraseology in sexual violence reporting, in which *alleged victims* are often juxtaposed with an accused who “strenuously denies” which suggests a much stronger position. A similar strong defense was found in the NexisSH news article corpus where the top ranking modifiers of the 2,840 instances of DENY (Table 69) include a variety of strong adjectives: *vehemently* (LogDice 11.5); *categorically* (LogDice 11.4); *emphatically* (LogDice 11.0); *strongly* (LogDice 10.8); and *flatly* (LogDice 10.1). A similar check on ALLEGE found no adjective modifiers within NexisSH to strengthen a statement. The alternative ACCUSE was also checked (Table 69) and the only adjective modifier found was *falsely* which would weaken rather than strengthen any claim. News publishers may be wary of being sued for defamation, and this would suggest that defendants may be more rich or powerful, giving them greater capacity to sue, than their accusers. This would align with news media power structures exercising the “neoliberal and patriarchal appeasement” described by Barca (2018, p. 280). The word *alleged* and the lack of equivalence between *alleged victim* and *alleged perpetrator* is discussed in section 2.5.8.

Search term	NexisSH
DENY	1,170
Top adjective modifier of DENY	<i>vehemently</i> 31 (11.5)
ALLEGE	1,647
Top adjective modifier of ALLEGE	0
ACCUSE	2,840
Top adjective modifier of ACCUSE	<i>falsely</i> 28 (LogDice 11.5)

Table 69, Frequency of ACCUSE and ALLEGE, along with their top adjective modifier.

### 7.3.6 Linguistic hi-jacking of VICTIM

Anderson (2020, p. 19) uses the term “linguistic hijacking” to describe how language is used to foster active ignorance and resist knowledge in order to “protect systems of domination”. For example the word VICTIM itself is used by detractors of the #MeToo movement to counter-accuse victims of sexual violence<sup>92</sup> using phrases such as “You are not the victim – I am/he is/she is!”. This use sets up the accusation as equivalent to the violence itself, and subverts the meaning so that the accused is repositioned as the real victim (Shoaf et al., 2024, p. 1910). An important aspect of social media is its potential to enable individuals who are subject to violence to build support communities. However, such resources are also accessible to critics who may use them to attack victims/survivors speaking publicly. When the accused suggests their voice as a victim is being silenced this may indicate neoliberal influences “where claims of being silenced flourish in the context of a growing entitlement to speak.” (Phipps, 2019, p. 70). In Example 235 a Democrat supporter admonishes *Bernie bros*<sup>93</sup>, for doubting Biden’s innocence. The writer then counter-accuses those who *believe #metoo bullshit*, instead proposing that Biden *is the victim*. The accusation is given additional negative prosody by denigrating the claim as *bullshit*, and the accuser as *that ugly woman*. Although it is not clear how the woman’s attractiveness is relevant, this may represent the “e-bile” used to silence women (Jane, 2014, pp. 531, 532).

<sup>92</sup> This reversal of an accusation is demonstrated in recent claims that the Black Lives Matter (BLM) movement is racist, usually accompanied by counter claims of either *All Lives Matter* or *White Lives Matter* in what is termed Linguistic Hijacking (Anderson, 2020, p. 2). However, the affirmative action in BLM is not racist because non-white populations are subject to systematic disadvantage that white people, the dominant group, are not. This incorrect use of words, whether *racist* or *victim*, by dominant groups to create resistance to the understanding of oppression, effectively becomes linguistic hijacking. Anderson’s main theme was racism, but included victimisation as an adjunct.

<sup>93</sup> Bernie (Sanders) and Biden were US Democrat presidential candidates.

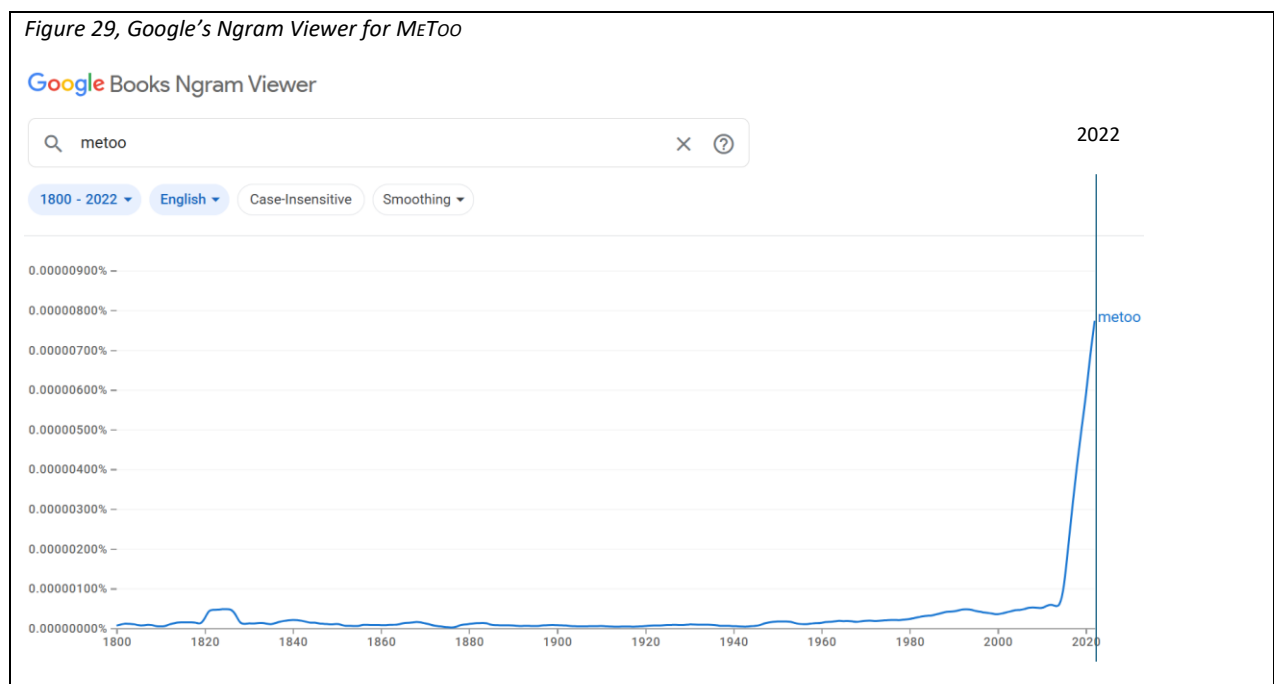
Example 235 Bernie bros still believe #metoo bullshit against Biden? Be smarter ok? He ,not that ugly woman, is the VICTIM .. ok?

(TwitterMT\_01to24, VICTIM, 2020-05)

This study found no evidence of SURVIVOR being hi-jacked in this way. An experimental substitution, based on Example 235, produces “He ,not that ugly woman, is the SURVIVOR .. ok?” which seems quite unlikely.

Used against any non-dominant group, narrative reversal attempts to neutralise the accusation and protect the system of domination. Additionally, by repeatedly associating VICTIM with false accusations, not only does VICTIM become associated with falsehoods in readers minds, but it is weakened as a legally significant word. The usefulness of the term victim with its long association with having been wronged and justice is damaged, along with its ability to challenge a systemic problem (Anderson, 2020, p. 15). Anderson (2020, p. 16) calls such shifts in meaning “enfranchised semantic drift” indicating that this kind of change tends to advantage the dominant group.

The consequences of linguistic hi-jacking can be demonstrated by an example of a reverse situation. Prior to the 1970s women were commonly subject to a range of unwelcome behaviours, often represented as humorous or flattering, illustrated in the 1952 cartoon in Figure 9, and for which there was no word to describe them as a wrong. The coining of the term *sexual harassment* as a key linguistic resource, enabled unacceptable behaviours to be recast as wrong. This freed a log-jam of accusations and discussion from the 1960s, shown in Figure 1. The initiation of the #MeToo movement in 2017 coincides with increased use of the term more widely in the texts referenced by Google’s Ngram Viewer (Figure 29). These examples illustrate the importance of a key linguistic resource to bring an issue into public awareness. The erosion or loss of an arguably even more important resource, VICTIM, should cause concern.



## 7.4 Pros and cons of SURVIVOR

The traditional understanding of SURVIVOR could be viewed as a sub category of VICTIM, but differentiated by being less needy and taking more personal responsibility (Cole, 2021, p. 30). In this view SURVIVOR retains some claim to justice, but avoids the full VICTIM stigma. While potentially advantageous for a specific SURVIVOR, active and public denial of VICTIM has the consequence of encouraging evaluation of personal traits, further denigrating the generalised VICTIM who, in contrast, is viewed as not taking personal responsibility and held to blame for their misfortune.

Termed as a SURVIVOR, the individual gains from the positive prosody around the term, but having to enact agency can be problematic if the individual has not truly reached that point of recovery (Sehgel, 2016). A second risk is that with the SURVIVOR represented as healed, the perception of harm recedes into the past, and the perpetrator's crime becomes more distant. Presenting as healed and restored, the individual may be expected to go public with a positive and empowering version of their experience for the benefit of others (Fohring, 2018, p. 160), discussed further in section 7.4.1. By providing a happy ending, both speaker and audience are subjected to less of a challenge to their belief in a just world. While some may find this a healing and empowering experience, it may be more risky for those not entirely healed (Fohring, 2018, pp. 154, 158). As increasing numbers of uplifting narratives of overcoming experiences of sexual violence and achieving good outcomes are produced, this becomes normalised and expected.

A word of warning comes from Lakoff and Bucholtz who showed how renaming an uncomfortable concept such as VICTIM may not solve the problem. They suggest that "Since attitudes toward the original referent are not altered by a change of name, the new name itself takes on the adverse connotations, and a new euphemism must be found" (Lakoff & Bucholtz, 2004, p. 51).

### 7.4.1 SURVIVOR as the only alternative to VICTIM

As discussed in section 2.5.5 there is a time line of healing that progresses from being a VICTIM to being a SURVIVOR of an incident. The duration of this healing, and any limit on eventual recovery, varies between individuals, and is influenced by a number of factors such as the nature of the event itself, but also the victim's personal resources and the practicalities of their environment (Dillenburg, 2008, p. 13). The healing process requires the individual to acknowledge their victimisation, and attempts to deny it only delays that process (Fohring, 2018, p. 160). The #MeToo movement encouraged all who had suffered sexual harassment to speak out, and the resulting shift in attitude has recast the identity of SURVIVOR as one marked by service to others and so obligated to speak out:

*"The perception that survivors of sexual violence must bear the burden to tell their stories is consistent with the implicit cultural assumption that victims are personally responsible for what happens to them—and what happens next."*  
(Delker et al., 2020, p. 4)

Delker is saying that it is the giving of public testimony that marks out a *survivor*.

Running a vignette study, Delker et al found that even when the culturally preferred redemptive story endings are provided (2020, pp. 3, 14, 15, 18) those suffering sexual violence were still viewed as stigmatised and less likable. Furthermore, having spoken out the speaker risks losing control of their story (Fohring & Horsfield, 2023, pp. 344, 345).

Returning to the current study, the modifiers of SURVIVOR and VICTIM, provided as Table 70 and Table 71, are very different. Moreover, when the same adjective appears for both VICTIM and SURVIVOR they are used for opposite effects. For example, adjectives such as *real* or *true* mark VICTIMS with positive attributes, and when contrast with VICTIMS not marked by such adjectives may delegitimise or suggest their claim is bogus. In contrast, *true* and *real* are used to emphasise the worthiness and legitimacy of the SURVIVOR, and not compared to another group. As VICTIM is more delegitimised, SURVIVOR not only avoids that stigma but is also elevated by the contrast, which suggests admirable traits beyond the average.

TwitterSH LogDice, (frequency)	TwitterMT LogDice, (frequency)	NexisSH LogDice, (frequency)	NexisMT LogDice, (frequency)
<i>real</i> (61, 9.64)	<i>real</i> (274, 11.26)	<i>alleged</i> (116, 10.92)	<i>alleged</i> (100, 11.23)
<i>actual</i> (42, 9.39)	<i>actual</i> (68, 9.67)	<i>actual</i> (5, 7.73)	<i>real</i> (15, 8.64)
<i>alleged</i> (23, 8.53)	<i>alleged</i> (57, 9.43)	<i>real</i> (6, 7.67)	<i>genuine</i> (5, 7.77)
<i>true</i> (15, 8.08)	<i>true</i> (37, 8.8)	<i>true</i> (4, 7.4)	<i>actual</i> (4, 7.37)
<i>genuine</i> (10, 7.61)	<i>genuine</i> (16, 7.7)		<i>perfect</i> (4, 7.37)
	<i>fake</i> (10, 6.77)		
	<i>legit</i> (6, 6.35)		
	<i>legitimate</i> (6, 6.34)		
	<i>innocent</i> (10, 6.9)		
	<i>perfect</i> (6, 6.29)		

Table 70, Delegitimising modifiers of the lemma VICTIM across the sexual violence corpora using WordSketch.

TwitterSH LogDice, (frequency)	TwitterMT LogDice, (frequency)	NexisSH LogDice, (frequency)	NexisMT LogDice, (frequency)
<i>actual</i> 6.55 (2)	<i>real</i> 8.51 (26)	<i>alleged</i> 6.06 (3)	N/A
<i>real</i> 6.43 (3)	<i>brave</i> 8.2 (11)		
	<i>true</i> 7.7 (9)		
	<i>actual</i> 7.52 (8)		

Table 71, Legitimising modifiers of the lemma SURVIVOR across the sexual violence corpora using WordSketch.

In Chapter 3 the power of Labelling Theory, either when assigned by others or as self-labelling, was described. The collocate *brave* in Table 71 for TwitterMT, is the only modifier of SURVIVOR, across all the corpora, that does not suggest veracity. *Brave* suggests that a person is “courageous, daring, intrepid”, and “stout-hearted” (Oxford English Dictionary [electronic resource], n.d.-h) which is quite a label to live up to. Bouvier (2022, p. 186) points out that such adjectives foreground an ideological evaluation of the social actor’s personal characteristics, distracting attention from what was done. The social good being called for in Example 236 requires the survivor to speak out publicly telling a story that is not just the simple truth as it might be for any other crime, but presenting it with an uplifting narrative and a positive and redemptive ending (Delker et al., 2020, p. 4). Cox (2020, p. 510) expands on this, suggesting that being positioned as a *brave survivor* also suggests someone who has overcome adversity, and “...typically involves a voluntary engagement with an acknowledged degree of personal risk to help others.”. In the context of this study, this criteria would be met by speaking out, which can provide support for others, and additionally, if a criminal case resulted in a conviction, then another unfortunate individual might be spared the experience of the same crime. Ironically, in order to pursue such a legal case against a perpetrator, the accuser must present as a *victim*. The risks incurred are significant, including scrutiny of their personal life, disbelief, threats, social media trolling, and being counter-sued for defamation. This puts a high cost on survivorship, and risking all of this would match the definition of *brave*. This would be a significant undertaking for anyone, but taken on prematurely could delay healing or bury it to resurface later on (Fohring, 2018, p. 160). Example 236 illustrates this obligation on survivors to speak out for the greater good.

Example 236 And to everyone saying ‘it should be handled privately’ the problem is that women aren’t believed and rape cases just aren’t taken seriously. That’s WHY as survivors we have to SPEAK OUT. It’s the only way to gain awareness and make our voices heard. #metoo” (TwitterMT, SPEAK OUT, 2019-09)

Fohring (2018, p. 162) suggests there are a cluster of behaviours, including weakness and vulnerability approximating to the Ideal Victim, that the victimised are expected to demonstrate. Contradicting this expectation, for example by being agentic and speaking out, is a violation punished with “victim-blaming, ostracization and exclusion”. In her earlier

research Fohring had discussed five behaviours that might reduce the harsh judgement associated with victimhood, all suggestive of the brave survivor trope:

*“...belittling the incident,  
envisaging a ‘worst world’,  
making downward comparisons,  
deriving benefit or learning from the experience, or  
considering themselves to have made an exceptional adjustment.”*

*(Fohring, 2012, p. 134)*

The most common reason that Fohring found for individuals deciding not to report was evaluating the act as “too minor”, and she suggests that this may be a consequence of minimising the event in order to avoid victimhood (2012, p. 158). In trying to convince others that the crime was minor, the speaker may also convince themself.

Writing in Harper’s Bazaar, Campoamor (2018, p. 3), a journalist and rape victim, highlights this forced heroism as misrepresenting reality, “silently promoting a super-human response that encourages victims to ‘get over’ an unspeakable violation.” In a similar vein, Sehgel (2016) writing in the New York Times, describes an expectation that individuals both provide personal testimony, and take responsibility for their own healing. A consequence being that visibility of both the harm, and the systemic causes of sexual violence crimes, are obscured. Sehgel is unconvinced by the voluntary engagement implied by the word *bravery*. She points out the low rates of reporting and convictions, and suggests that the victimised are *resilient* rather than *brave* because they have little alternative.

Returning to the corpora, Table 72 shows the results of WordSketches of BRAVE within the four sexual violence corpora. *Survivor* in all four has a stronger collocation with BRAVE than does *victim*, excepting NexisSH\_01to24 which has no collocations for either.

	TwitterSH LogDice, (frequency)	TwitterMT LogDice, (frequency)	NexisSH LogDice, (frequency)	NexisMT LogDice, (frequency)
BRAVE	<i>victim</i> 5.3 (2) <i>survivor</i> 8.2 (3)	<i>victim</i> 5.7 (4) <i>survivor</i> 8.2 (11)	<i>victim</i> 0 (0) <i>survivor</i> 0 (0)	<i>victim</i> 5.5 (1) <i>survivor</i> 8.5 (4)

Table 72, Modifiers *victim* and *survivor* of the lemma BRAVE using WordSketch.

Looking more closely at the BRAVE concordances, the writer of Example 237 acknowledges and heaps praise on the SURVIVORS who *courageously came forward*, while in the present tense *will not be silenced* sets up an expectation of their stamina. Even such support may present a conflict. Having been publicly lauded as *brave* and *courageous* may make any later loss of nerve and retraction difficult without losing face. Failure to continue *courageously forward* may feel like failing to meet the expectations created, and letting down the other *brave survivors* inferred by the hashtags.

Example 237 Women who have endured sexual harassment and abuse will not be silenced. Thank you to the BRAVE SURVIVORS who courageously came forward.  
#TimesUp #MeToo #BeHeard #StandWithWomen

(TwitterSH, BRAVE SURVIVOR, 2020-03)

The dichotomy between the obligation on an individual to speak out and its cost, is expressed clearly in Example 238 as a *worst nightmare*, with only two equally unpleasant options. This kind of dilemma is unusual in crimes outside of sexual violence. For instance a burgled householder is unlikely to be burdened by guilt if they did not publicly tell all.

Example 238 My second worst nightmare is being forced to share with the world that i am a SURVIVOR of domestic violence my first worst nightmare is not telling anyone and knowing that I could have helped even just one person by sharing my story.  
MailOnline (NexisMT, SURVIVOR, 2020-12)

Example 239 provides an interesting example, contrasting the *courage* needed to articulate victimisation, against others who give in to *fear*, emphasised by their *choose[ing] to stay silent*.

Example 239 Do you know why sexual harassment, physical abuse, and bullying are still a prominent crisis? Because some of the VICTIMS choose to stay silent because of fear. SO NEVER EVER DEGRADE SOMEONE WHO CHOOSES TO SPEAK UP. YOU DONT KNOW HOW MUCH COURAGE IT TAKES.

(TwitterSH, VICTIM, 2020-11)

While the writer of Example 239 defends victims, they also implicitly blame them for the *crisis* rather than the harassers, abusers, and bullies, or the poor experience of the criminal and justice systems, reinforcing responsibility as personal rather than political.

This suggests a wider environment where JWT is operating to keep the wider population feeling safe in the knowledge that provided they do the “right” things, they will remain safe.

#### **7.4.1.1 Time and healing**

The criminologist Ben-David (2020, pp. 10, 12) describes an extended process of healing from *victim*, through *survivor*, to *thrivers*<sup>94</sup>, and finally to *overcomer*, a process which is neither simple nor painless. She states that most take months or years to reach the *survivor* stage of recovery, and it can take a lifetime to become an *overcomer* (Ben-David, 2020, pp. 7, 12). Experiencing guilt and self-blame are part of this process, and any pressure on the individual to deny victimisation may be harmful (Ben-David, 2020, p. 11). While getting stuck at the *victim* stage is a potential problem, so is a premature adoption of a *survivor* identity, which would cause an individual to conceal vulnerability and feign resilience that prevents them from fully utilising support (Ben-David, 2020, p. 12).

This healing sequence is illustrated by the following set of unrelated tweets from Example 240 to Example 242, all of which are personal, first person disclosures. This also illustrates FPDA’s notion that the degree of power an individual has shifts with time and situation (Baxter, 2003).

---

<sup>94</sup> The terms *thrivers* and *overcomer* are not used in the present study.

The child writing in Example 240 is clear that she is a victim and is fully in the grip of trauma from her experience so that she *can't even think straight any more*.

Example 240 reminds me of him. i cant even think straight anymore. my name is sarah. im fourteen, and I AM A VICTIM of sexual assault. #MeToo  
(TwitterMT, I AM A VICTIM, 2020-12)

The writer of Example 241 alludes to some degree of healing in that only *sometimes in crowded places I can be anxious...* implying this was once usually the case, but does not always happen now. *But to be honest* suggests that the candid disclosure *There is still a bit left*, is not normally publicised and is now shared in confidence. This may indicate a concern that the *social phobia* might be judged as a moral failing, and not entirely acceptable. The final sentence is slightly ambiguous, but could be read as *I am still overcome[ing] it*, indicating the ongoing process of healing.

Example 241 And I do develop social phobia before... But now I have mostly overcome it ???????? But to be honest, There is still a bit left.. And I AM A VICTIM of sexual harassment.. So sometimes in crowded places I can be anxious... And I am still overcome it ????  
(TwitterSH, I AM A VICTIM, 2020-01)

In Example 242 the writer describes a later stage of healing where they believe that they can *joke around* and *hide my pain well*. The fact that there is pain to hide suggests that healing is not complete, but it appears the individual has reached a stage where they have enough capacity to consider others' needs. The unspoken assumption that as a *survivor* the individual will have the resources to speak out, underpins the obligatory part of survivorship. In this case however, *sometimes I feel like opening up about what happened & show others they are not alone*, suggests an authentic and spontaneous altruistic desire rather than obligation.

Example 242 I tell myself: 'Someday I won't think about what happened on November 14<sup>th</sup>' but every year I think about it. I joke around a lot & I hide my pain well, but sometimes I feel like opening up about what happened & show others they are not alone. I AM A SURVIVOR. (TwitterMT, I AM A SURVIVOR, 2020-11)

The complete transition from VICTIM to SURVIVOR is summarised in Example 243, from the past tense *I was a VICTIM*, by way of the present *now I am a survivor*, to the future *i will never be what happened to me*. However, the phrase *i will never let it define me as a person*, suggests an element of choice which perhaps only comes when healing is complete. There is also an element of defiance in this tweet suggesting that this may be a statement of intent as much as a statement of fact.

Example 243 six years ago today i was a victim. now I AM A SURVIVOR . i will never be what happened to me on this day; i will never let it define me as a person.  
#MeToo (TwitterMT, I AM A SURVIVOR, 2020-12)

While potentially an empowering position for the writer, such statements further diminish VICTIM for those who have not healed, and who may lack the personal and external resources to do so. Example 243 conveys an implicit suggestion that a victim would *let it* define them, which implies that the victim has the capacity to prevent it. With increasing toxicity of VICTIM,

pressure may increase for others to opt for a SURVIVOR identity prematurely, and as more individuals distance themselves from VICTIM, it becomes more tarnished.

SURVIVOR can provide advantages over VICTIM for the individual, however these are not guaranteed or universal, and their use within the specific context should be conscious, considered, and respectful (Papendick et al., 2017, p. 19 (in pdf reader)).

## **7.5 Who is missing from pay gap discussions?**

This section discusses the individuals who are disadvantaged by the pay gap but not referenced in the text, and this requires the field of vision to expand to the whole corpora. In both the TwitterPG and NexisPG corpora, individuals negatively impacted by the pay gap are frequently Backgrounded while discussion focuses on the phenomenon of the *gender pay gap*. Within this study the term *gender* means women, and is used in preference to *sex* because the pay gap and its causes are viewed as socially constructed. Where they are referred to, the disadvantaged are most frequently labelled as *employee(s)* or *women*. However, neither *women* nor *employees* differentiates between those affected by the pay gap and those unaffected, nor do they carry the association of a wrong in the same way that *victim* or *survivor* does.

Many employers and employees will be unaware of a pay gap, or even an equal pay issue because of unconscious gendered bias. RET (Wagner & Berger, 1997, pp. 9, 12), discussed in section 2.4.5, suggests that those who are paid more or viewed as higher status are judged as being more skilled and higher performing, resulting in both women and men appraising men's work as more valuable. This contributes to the glass ceiling issues of recruitment, pay, and promotion.

National level gender pay gap calculations include non-working parents and carers which are predominately mothers. While this might be a free choice for some women, for others it might be constrained by structural barriers such as lack of affordable childcare. Nardone (2018, p. 183) was able to identify such social problems in her Italian language corpus, however for the present study this is not possible, so full time unpaid parents and carers are missing.

The missing individuals discussed in the following sections qualify as missing, either because they are not acknowledged in the discussion at all, or if they are acknowledged their disadvantage is not.

The next section, 7.5.1, uses a high level corpora search to look for traces of the missing. Section 7.5.2, identifies all of the known demographics that can be removed from the search. Then the specific issues obscuring those in low paid and career segregated roles are outlined in section 7.5.3.

### **7.5.1 Corpora level**

A high level corpora interrogation provided some outline information. TwitterPG contains 75,000 tweets with 22,983 instances of WOMAN. If WOMAN occurred only once in a tweet, there would be 52,000 tweets that did not include WOMAN. TwitterPG also contains 1,808 instances of EMPLOYEE, so if EMPLOYEE occurred only once in a tweet, there would be 73,000 that did not mention EMPLOYEE. While tweets are often fragments of conversation which may

have mentioned those disadvantaged elsewhere, this still suggests a significant lack of visibility.

A similar exercise on NexisPG reveals that its 1,898 news articles contain 21,056 instances of *woman* and 2,927 of *employee*. If these were evenly distributed then each article would contain 11 instances of *woman*, and 1.5 instances of *employee*. Even allowing for the smaller numbers of other descriptors such as *doctor* or *teacher*, this suggests that TwitterPG has greater numbers of the missing than NexisPG.

When the pay gap is dismissed as bogus a further group of the disadvantaged are removed, because if the pay gap is non-existent, then no one is disadvantaged by it. Table 62, included 1,015 instances of negating adjectives in TwitterPG that represent the pay gap as bogus, and all of the individuals discredited by these words are distanced from the discussion.

### 7.5.2 A process of elimination

Although we cannot directly specify who the missing are, we can narrow down the field by identifying and eliminating who they are not. A WordSketch of TwitterPG nouns that indicate professions modifying FEMALE returns 60 job titles, those with a frequency of 10 or more are shown in Table 73, and the full list is given in Appendix E.

Profession	Freq	LogDice	Profession	Freq	LogDice
<i>doctor</i>	52	9.6	<i>sport</i>	15	7.8
<i>athlete</i>	42	9.4	<i>nurse</i>	14	7.8
<i>professor</i>	36	9.2	<i>surgeon</i>	13	7.8
<i>player (soccer)</i>	44	9.1	<i>executive</i>	13	7.7
<i>leader</i>	31	8.9	<i>artist</i>	12	7.7
<i>CEO</i>	31	8.8	<i>actor</i>	12	7.6
<i>physician</i>	25	8.7	<i>MP</i>	11	7.5
<i>model</i>	22	8.4	<i>writer</i>	11	7.5
<i>graduate</i>	19	8.2	<i>teacher</i>	12	7.4
<i>firefighter</i>	17	8.2	<i>freelancer</i>	10	7.4
<i>footballer</i>	15	8	<i>entrepreneur</i>	10	7.4
<i>engineer</i>	15	8	<i>founder</i>	10	7.4

Table 73, Job title nouns modifying FEMALE (total in table 492)  
Within TwitterPG using Word Sketch of, with a frequency of 10 or more.

Sports people's salaries and prize money are usually in the public domain, as evidenced by the huge interest in the rewards of the US women's football team following their World Cup win (Alkhamash, 2021). The salary bands of state employed teachers and medical professionals are also usually published. Additionally, listed companies with more than 250 employees need to declare CEO salaries in the UK (since 2018), and in the US (since 2007) which often prompts public discussion of executive pay. This supports the possibility that transparency may play a part in identifying the disadvantaged.

Turning to NexisPG, profession nouns that modify FEMALE with a frequency of 10 or more are given in Table 74.

## Sexual violence and the gender pay gap

### 7 The labelled and the missing

Profession	Freq	LogDice	Profession	Freq	LogDice
<i>graduate</i>	90	10.0	<i>director</i>	27	8.0
<i>doctor</i>	55	9.4	<i>physician</i>	19	7.9
<i>candidate</i>	42	8.9	<i>professor</i>	19	7.9
<i>CEO</i>	39	8.9	<i>employment</i>	18	7.8
<i>leader</i>	39	8.7	<i>lawyer</i>	17	7.8
<i>player</i>	32	8.6	<i>staffer</i>	16	7.7
<i>executive</i>	34	8.4	<i>boss</i>	16	7.6
<i>presenter</i>	27	8.4	<i>actor</i>	15	7.6
<i>athlete</i>	25	8.4	<i>minister</i>	17	7.5
<i>star</i>	27	8.3	<i>owner</i>	12	7.2
<i>manager</i>	28	8.3	<i>artist</i>	11	7.2
<i>barrister</i>	22	8.2	<i>engineer</i>	11	7.2
<i>entrepreneur</i>	22	8.2	<i>MP</i>	12	7.2
<i>partner</i>	23	8.1	<i>surgeon</i>	10	7.0

Table 74, Job title nouns modifying FEMALE (total in table 725)

Within NexisPG using Word Sketch of, with a frequency of 10 or more.

Each of the profession nouns was categorised according to the Industry categories in Figure 30, which shows the distribution of the debate across industries within the pay gap corpora (Table 75). Some industry categories are very broad, such as Arts, entertainment & recreation which includes cricketer, celebrity, and pornstar. Additionally another category was created, highlighted with an \*, called \*Executive & Managerial for the large number of executive and managerial professions listed that could not be aligned to an industry from Figure 30.

While these categories are not a perfect fit for the data of this study, they do highlight that some industries are more frequently discussed, with \*Executive & Managerial (367) the most frequent overall and for NexisPG (262), and the Arts, entertainment & recreation category having the highest frequency for TwitterPG (211).

The full lists of categorised profession nouns are given in Appendix E.

	TwitterPG (frequency)	NexisPG (frequency)	Total
*Executive & Managerial	105	262	367
Arts, entertainment & recreation	211	137	348
Health & social work	119	103	222
Other services	39	140	179
Professional, scientific & technical	48	58	106
Public administration & defence	30	63	93
Information & Communication	38	38	76
Education	30	34	64
Transport & storage	8	6	14
Administrative & support services	3	0	3
Accommodation & food services	3	0	3
Financial & insurance	0	0	0
Construction	0	0	0
Manufacturing	0	0	0
Wholesale & retail	0	0	0
Real estate	0	0	0

Table 75, Categories of profession nouns modifying FEMALE

Low frequencies of profession nouns within an industry may arise for various reasons, for example a heavily male dominated industry with few women to challenge and gain media or public attention. Alternatively, within roles and industries with equal pay, women may be unaware of the wider causes of the gender pay gap. A surprisingly visible role within the TwitterPG corpus is RECEPTIONIST, although it does not appear as a WordSketch modifier of FEMALE, and *female receptionist* does not exist within the corpora. However, *receptionist* does occur 47 times (14.85 pmw) within TwitterPG, and is always used as an example of a nonsensical contrast with a heavily segregated male role. In fact the top collocate of RECEPTIONIST is *firefighter* (LogDice 10.52), and *female firefighter* is used to highlight that women do not commonly choose this career, illustrated in Example 244. Despite the popularity of this type of argument, neither the pay gap nor equal pay definitions suggest that entirely different occupations should receive the same pay.

Example 244 Do t get me wrong gender pay gap is real and wrong, but J don't think comparing a firefighter to a RECEPTIONIST is a good comparison... running into a burning building vs sitting at a desk answering phone calls  
(TwitterPG, RECEPTIONIST, 2019-06)

While RECEPTIONIST is visible in the discourse, receptionist's disadvantage is more likely due to career segregation rather than an equal pay issue. As their discrimination is not recognised, they join the ranks of the missing. In contrast, RECEPTIONIST appears only four times in NexisPG, and never collocates with *fire fighters*. Career segregation is discussed in the next section.

### 7.5.3 Career segregation

Despite equal pay being written into law in the US (1963) and the UK (1970), obtaining equal pay for similar work is still an issue for some. This is exacerbated by the onus being on individuals who are discriminated against to raise and fight the case. Additionally, the financial risk of losing, and having to pay the employer's legal costs, may be a disincentive for many employees (Citizen's Advice Bureau, n.d.; Cooke, 2024, p. i357). Consequently, some women may be aware that they are disadvantaged, but lack the funds to take legal action, and so are missing. The rest of this section looks beyond the corpora to external statistical sources.

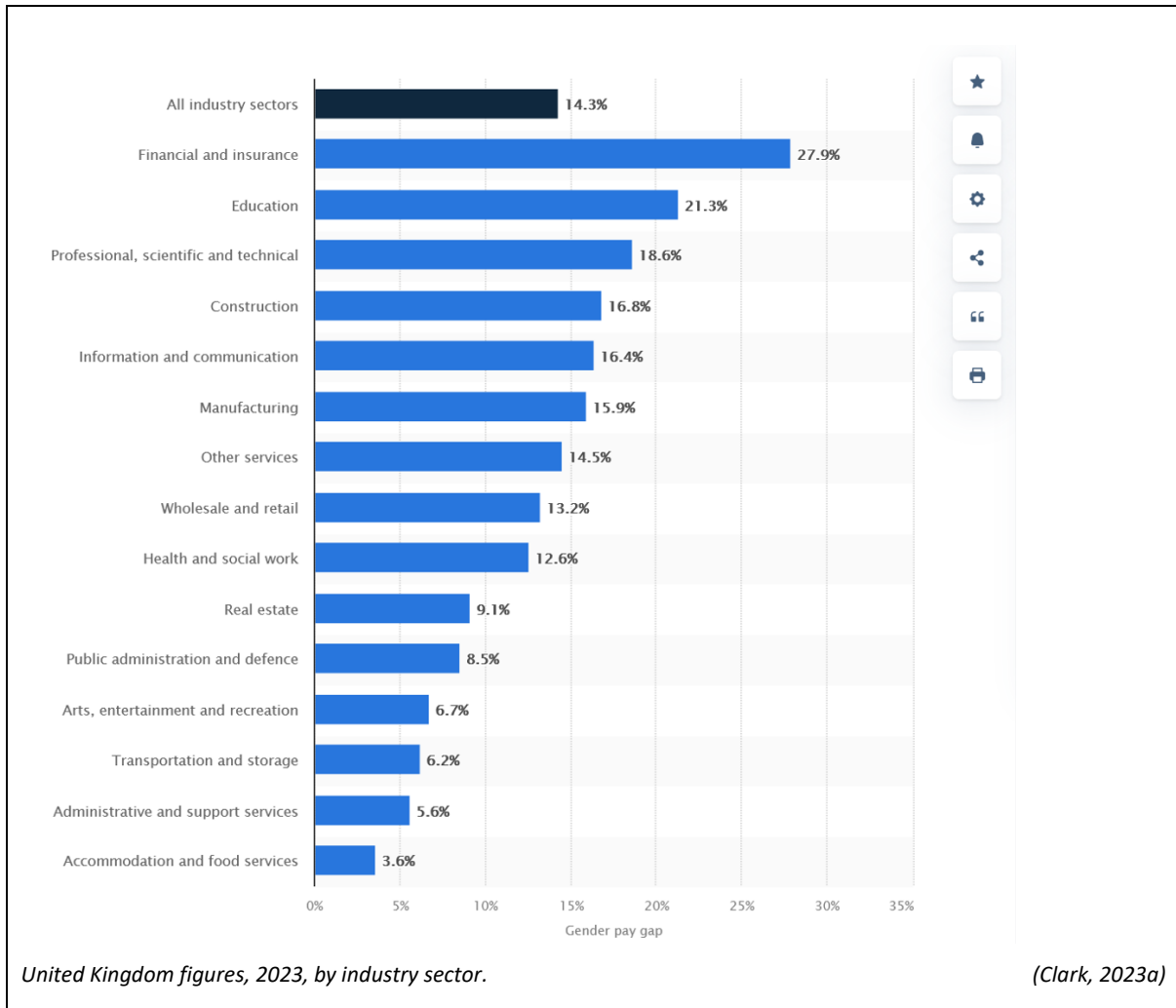
The greatest pay gaps, at the top of the table in Figure 30 (Office for National Statistics, 2023) include the higher paid industry sectors and Executive & Managerial, discussed in section 7.5.2. These professional roles are associated with higher levels of education, and greater expectations of equal pay and opportunity, so disadvantaged women in these sectors are less likely to remain silent and missing.

In the corpora of this study there was little evidence of discussion of remuneration in traditionally lower paid roles, with *cleaner* (3) being the only example on the tables of Appendix E, suggesting their wages are not reported in news articles or discussed on Twitter. These employees may lack the information necessary to identify unfair hiring or promotion practices, or even basic equal pay disparity.

In low paid jobs the hourly pay difference between male and female employees is smaller because minimum wage legislation creates a floor (Platt, 2024, p. 291). Women are overrepresented in healthcare and education that often have lower pay (World Economic Forum, 2025, pp. 49, 50) and fill more part-time jobs which have comparatively lower median pay (Office for National Statistics, 2023, pp. 4, 7). Additionally, more women than men earn less than the living wage, with 59.5% of low paid jobs are held by women (Living Wage Foundation, 2024).

The sectors at the bottom of the chart in Figure 30 (Clark, 2023a) are likely to be dominated by female retail assistants, waitresses, and cleaners (Living Wage Foundation, 2024). These women may have equal pay to their male colleagues, but some are likely to be subject to career segregation, and others juggling unpaid care work. They may be ignorant victims, like Christie's disadvantaged school child, unaware of their disadvantage, discussed in section 2.5.1.1. Without an identifiable perpetrator, or specific crime, making a complaint becomes difficult. Those with few resources and low expectations of justice may decide that they have little to gain by taking action, and so remain missing.

Figure 30, Gender pay gap for gross hourly earnings



When low paid groups are visible, it may be due to strong union representation, as discussed in Example 245 which refers to the Birmingham (UK) City Council scandal raised in 2012 by the GMB union, and not settled until 2024. Unionisation is associated with smaller pay gaps and higher wages (Gould et al., 2016, p. 23). The term “*missed out*” (Oxford English Dictionary [electronic resource], n.d.-p) suggests their payment was accidentally overlooked, rather than deliberately excluded. The use of quotation marks may suggest that the writer is drawing attention to this distinction.

Example 245 In the UK, women choosing traditionally female roles only account for 22% of the gender PAY GAP, compared to a significant 37% resulting from direct discrimination. Eg, Birmingham City Council paid £757m to settle equal pay claims brought by women who “*missed out*” on bonuses.

(TwitterPG, PAY GAP, 2019-04)

This suggests that it is more likely that missing individuals are not union members.

## Sexual violence and the gender pay gap

### 7 The labelled and the missing

In 2023 60% of low paid jobs were held by women (Living Wage Foundation, 2024, p. 1), and according to Statista (Clark, 2023b), women have always held more than half of zero-hours contracts, 58% in 2023 (Figure 31).

Figure 31, Gender split of zero-hours contracts



The voluntary Real Living Wage is higher than the Statutory Minimum Wage because it reflects the true cost of living. In 2023 the UK sectors with most employees earning below the Real Living Wage (£10.90) included: Accommodation and Food Services (48.1%), Wholesale and Retail (23.2%), and Administrative and Support Services (19.9%), shown in Figure 30. Many of the women in these sectors will have equal pay with their male colleagues, but the issue is in their over-representation in these low paid jobs. Subject to gendered expectations, opportunities, and social barriers that restrict their options, they remain invisible.

In summary, the missing women disadvantaged by the pay gap are likely to be predominantly:

- Working within industry sectors that lack pay transparency;
- Working in career segregated roles;
- Working in low pay roles;

- Working in non-unionised environments; and/or
- Having low personal expectations of justice.

## 7.6 Who is missing from sexual harassment discussions?

Some of the types of silencing discussed in relation to the gender pay gap also apply to those missing from sexual violence discourse. However, the victimised within sexual violence discourses are most commonly characterised in two ways. Firstly, the visible victimised individuals are more often delegitimised; and secondly, there is an un-named and silent host of victims marked with legitimising adjectives who are used as a contrast to a specific bogus victim. Hence, the former may lose their victim status by being discredited as not being victims, while the latter do qualify as victims but being anonymous cannot communicate the crime.

The victims/survivors appearing in the corpora of this study represent the tip of an iceberg that has started to emerge over the last 50 years, starting with the coining of the phrase *sexual harassment*. References to the #MeToo movement increased from 2017<sup>95</sup> illustrated in Figure 29.

Some people have been shocked by the high numbers of sexual violence accusations and disclosures of the #MeToo movement, but statistics indicate that 99.4% of incidents still go unreported (Tranchese, 2024, p. 4). She quantifies the problem: fewer than one in six rape or assault victims reported the crime to the police; 2.9% of recorded rapes resulted in prosecution; and just 0.6% of reported rapes are false allegations (UK Crown Prosecution Service figures). This indicates that there remains a large number of individuals who are missing from official data, leaving society's understanding of sexual violence as partial, and unbudgeted for by the police, judiciary and support services. Many of these missing victims/survivors are likely to also be missing from the discourses of the sexual violence corpora.

Reported crimes featuring in news media are often positioned as isolated episodes, emphasising their novelty and newsworthiness. This distracts from the reality of the frequency of common patterns of events, underplaying the magnitude of the problem, and obfuscating ideological links to systems such as neoliberalism (Barca, 2018)

The ONS Crime Survey for England and Wales (CSEW) noted that victims of sexual violence were likely to under-report even to their anonymous survey (2025-b, p. 25 (in pdf reader)). Some do not speak out because they do not view themselves as subject to a crime. In the first instance, the individual must recognise the incident as a crime. A study of Danish workplaces found that years of silence can result from: ignorance of what sexual harassment is; the reaction of others in normalising sexual harassment; and use of a gradual grooming process by a perpetrator (Nielsen et al., 2024, pp. 2, 4, 6, 7). Even when aware of their victimisation, Nielsen's participants were still likely to remain silent fearing retaliation and career damage.

---

<sup>95</sup> The #MeToo movement was initiated in 2017 by Alyssia Milano.

As victimisation and crime are associated with a victim, individuals may have to deny the crime in order to support their denial of the stigmatised victimhood, and restricting their access to justice and support services.

There are numerous examples of concordances in earlier chapters of a third party referring to marked *real* and *true* victims who are used as comparators, and against whom a specific individual falls short. The marked are usually represented as a generalised, unspecific, and almost abstract group who are nevertheless worthy and deserving. While they are deserving of support and justice, this group is not sufficiently tangible to make any demands on society, or threaten any individual. Thus anonymous *real* and *true* victims disappear, joining the ranks of the missing.

Not all who are missing now will remain silent. Some might gauge the current risks of speaking out to be too great. For example, reporting a manager for rape might risk disbelief or condemnation by others, or retaliation if they remain in place. With time and repetition, similar complaints may be taken more seriously, and if the victim then feels sufficiently healed, they may then speak out and report. One of Nielsen's (2024, p. 8) participants described how remaining silent initially provided protection, "The more I talk about it, the more real it gets. When I was just thinking about it, I could pretend it never happened." Recent understanding acknowledges that victims cannot always report immediately, and a period of time and healing is needed before they can address the situation. This, however, is not without consequences, as circumstances that lead to delayed reporting and access to support and justice services may impact victim/survivor's mental health (Delker et al., 2020, p. 16).

### **7.6.1 Who benefits from silence?**

Some who do report find that the systemic barriers exhaust their personal and external resources, preventing them from prosecuting legal cases to a conclusion (Delker et al., 2020, p. 17). Those barriers thus benefit perpetrators by eliminating complaints, prosecutions, and convictions. Tweets may position accusations that are withdrawn, or result in any other conclusion than a guilty verdict as proof of a fraudulent claim. Tanushree Dutta's sexual harassment court case against Nana Patekar illustrates this. In 2019 the court closed the case (BBC, 2019) finding insufficient evidence, however this was interpreted by some as evidence of malicious defamation. Example 246 compares the guilty verdict of Harvey Weinstein, with the insufficient evidence finding for Nana Patekar and the reaction of local police who make unsubstantiated claims against Dutta.

Example 246 USA: Harvey Weinstein sentenced to 23 years. In India : Patekar was cleared of the sexual harassment charges by the police. The report filed by the Oshiwara police in Mumbai said that the complaint filed by Tanushree could be "malicious" & "out of revenge". #MeToo #metoindia (TwitterSH, 2020-03)

In Example 247 the writer goes further calling for punishment of those raising false claims, suggesting this writer also interprets insufficient evidence as proof of defamation.

Example 247 No evidence found against Nana Patekar .... will Law punish those...  
Who falsely implicates men in Rape cases, Sexual harassment  
cases????????????? <https://t.co/3HZ9EmGCjO> (TwitterSH, 2019-05)

The counter accusation of those speaking out as harming others was discussed in section 4.5.2.

### **7.6.2 The dangers of speaking out**

Within some cultures victims will find it almost impossible to safely report crime and obtain justice. A common cross-cultural thread is that a victim is likely to get the best results when they adhere to local expectations of victim behaviour (Dunn, 2008, p. 5 (in pdf reader)). Before speaking out victims are likely to carefully weigh the pros and cons of reporting, which can be devastating. Equality Now (2017, p. 6) named nine countries where a perpetrator can avoid prosecution for rape if they marry the victim. Victims may prefer to remain silent, pragmatically they may understand that the costs of reporting would outweigh the benefits (Cooke, 2024, p. 358).

## **7.7 Conclusion**

This chapter addressed how the labelling as a SURVIVOR or a VICTIM affects individuals, and additionally what this means for those who are missing from the discourse and therefore unlabeled and silent.

RQ4A asks if the SURVIVOR label is more beneficial than the VICTIM label for the victimised, and for society? The answer is that it depends. Under the right circumstances SURVIVOR can provide a positive and beneficial identity for individuals that VICTIM does not. In other situations SURVIVOR can be a heavy burden and even delay healing. At societal level, displacement of the term VICTIM by SURVIVOR reduces the sense of wrong-doing and distances the perpetrator. Furthermore, the growing preference for SURVIVOR may reflect a right-leaning culture and its preference for self-help solutions and preservation of JWT beliefs.

In adopting the label SURVIVOR, the individual avoids the stigmatisation of VICTIM, and gains social approval for taking responsibility for overcoming their experience. Thus SURVIVOR provides an attractive and strong image, and may seem the obvious choice between the two available alternatives. However, this choice is not made in a vacuum but results from the long-running denigration of VICTIM that has experienced increasing negative prosody until it is used as an insult. It is concerning that the victimised may be unaware of the hidden costs of adopting a SURVIVOR identity, which include obligations and expectations that may be challenging. Part of the obligation of those choosing SURVIVOR, in the West at least, is to present their story as uplifting with a redemptive ending. This converts an “unspeakable violation” to something less likely to unsettle the audience’s belief in a just and comfortable world (Campoamor, 2018, p. 2). In reality, any choice is between poor alternatives and often appears to be about avoiding VICTIM rather than choosing SURVIVOR because “There is simply no alternative.” (Sehgel, 2016, p. 2 (in pdf reader)). For society there is a risk that the denigration and stigmatisation of VICTIM erodes its usefulness as a key term for calling out injustice. VICTIM remains the term used by the legal system, and allows discussion of victimisation and victimisers, foregrounding the wrong-doing and perpetrator in a way that SURVIVOR cannot.

RQ4B concerns the individuals missing from the sexual violence and pay gap concordances; that is those who might have reasonably been expected to be there. For the sexual violence corpora, the figures for under-reporting indicate that the reported cases represent the tip of an iceberg, discussed in section 7.6. Of those who do report, some are represented as a bogus un-marked VICTIM, or an anonymous but deserving *real victims*.

For the pay gap corpora, discussed in section 7.5, the missing are more loosely defined by who they are not. Eliminating the visible leaves the workers who are more likely to be: non-unionised; in traditionally female roles; working part-time and/or zero-hours contracts; in sectors lacking pay transparency; and who have low personal expectations of justice. Many of those workers may be unaware of the discrimination. Unconscious gender bias of both employers and employees contributes to invisibility within the pay gap, discussed in section 2.4.5.

Self-evidently, the missing are silenced and do not get to choose their labels, and are not given one. Voiceless, they are unlikely to obtain support or justice, and the unchallenged perpetrators continue to perpetuate the problems within the systems that maintain the status quo.

## 8 Conclusion

There has been much progress on removing barriers to women's full participation within the workplace, and society has become less tolerant of sexual violence. However, the more recent shift in developed nations towards a more populist right has given rise to a loss of a number of important women's rights, particularly in the US (section 2.3.2). This suggests that there is a current need for research into the role of language in both driving social change, and as a means of identifying and measuring shifts in society. A Lancaster library search (section 1.2) indicated that research combining SEXUAL HARASSMENT, AND PAY GAP from a linguistic perspective was sparse, and this study was conducted with the intention of contributing to the understanding of this area.

In the first half of this final Chapter (section 8.1) I summarise the key findings of this study and how they relate to the research questions. Then in the second half I reflect on the experience of studying this topic (section 8.2). This is followed by my thoughts on the limitations of this study (section 8.3) and then some suggestions for future research (section 8.4). The contribution of this study and its implications is discussed (8.5). Finally, I draw together the threads of the analysis with some concluding remarks (section 8.6).

### 8.1 The key findings

The overarching research question driving this study asked how those harmed by gender-based discrimination were represented in the discourse, how this relates to the ideological landscape, and who this benefits. This is broken down into the four research questions (sections 8.1.1 to 8.1.4). Each section starts with a reminder of the research question before setting out my response.

#### 8.1.1 RQ1

Research question 1 focusses on the sexual violence vL500 datasets, which I address in three parts. The analysis for this question uses the manually coded +/- prosody flags (Chapter 4).

RQ1 Within the context of Twitter and Nexis sexual violence datasets, how are VICTIM and SURVIVOR represented in relation to possible positive<sup>96</sup> and negative affective responses? This question is addressed in three parts:

RQ1A, Are VICTIM representations more frequently negative than SURVIVOR representations? (sections 4.2 & 4.3)

RQ1B, What are the dominant themes associated with VICTIM and SURVIVOR? (sections 4.4 & 4.5)

RQ1C, What is gained by claiming the SURVIVOR label? (section 4.6)

The answer to RQ1A is yes, VICTIM representations are more frequently negative than SURVIVOR representations.

The analysis for RQ1A shows that while the majority of concordances within all datasets are coded as having positive prosody I found significantly more negative representations for

---

<sup>96</sup> Positive and negative is defined in section 3.4.2

VICTIM than for SURVIVOR (sections 4.2). Nexis datasets had no negative representations of SURVIVOR, while Twitter datasets had three times as many negative representations for VICTIM compared to SURVIVOR. The vLTMT500<sup>97</sup> dataset was the most strongly negative with 18% of its representations for VICTIM coded negative, compared to 5% for SURVIVOR. Hence, the highest frequencies of negative representations are associated with the intersection of the VICTIM lemma, the Twitter corpora, and the #MeToo search term.

The answer to RQ1B, is that five themes emerged that mapped directly to van Leeuwen social actor types: Activation, Appraisal; Beneficialisation; Indifferentiation; and Subjection (sections 4.4 and 4.5). Additionally I found five subthemes emerged within Activation: Blaming/shaming, Consequences, Playing victim, Speaking out, and Victim vs survivor. Many apparently negative themes, such as Playing victim were inversions, where a phrase from an earlier negative tweet was being called out, resulting in an overall positive prosody coding. Conversely, some codings which suggested positive personal agency, were later reframed as Blaming/shaming because an action presented the victim/survivor as causing their own victimisation.

I found representations of SURVIVOR were associated with positive prosody more frequently than VICTIM for all themes. Positive representations contained a predominance of advocates speaking out against unacceptable behaviours. Concordances coded with negative prosody show all of the same themes as the positive concordances, but they varied in frequency. VICTIM appears at frequencies at least three times higher than SURVIVOR for all themes bar Indifferentiation. Expressing solidarity, Indifferentiation is three times more frequent for SURVIVOR than VICTIM. My analysis of the themes confirms that VICTIM is represented more negatively than SURVIVOR.

The answer to RQ1C is that in opting for the SURVIVOR label an individual is more likely to gain a positive and agentic representation, less likely to be blamed or shamed, and less likely to be accused of deceit. Some individuals within these datasets claim personal benefit from the stronger and more agentic identity of SURVIVOR. The stronger association with positive prosody suggests the SURVIVOR concordances may be the culturally preferred term. However, agency comes with an expectation of personal responsibility that may have negative consequences (Chapter 7). The changing dictionary definitions appear to be in line with the language use seen in the study's corpora (Chapter 1).

---

<sup>97</sup> The naming scheme for corpora and datasets is provided in Appendix B.

### 8.1.2 RQ2

The second research question addresses the pay gap (Chapter 5). This question is based on the pay gap vL500 datasets.

RQ2 How are those disadvantaged by the gender pay gap represented within tweets in relation to positive and negative affective responses, and how does this compare to representations associated with sexual harassment? This question is answered in three parts:

RQ2A Do the pay gap datasets contain more negative representations than the sexual violence datasets? (section 5.3)

RQ2B What are the dominant themes associated with those disadvantaged by the pay gap? (section 5.4 & 5.5)

RQ2C How are the themes found in the pay gap datasets different from those of the sexual violence datasets? (sections 5.5 & 5.7)

The answer to RQ2A is yes and no, because vLTPG500 (21%) and vLNPG500 (0%) are the most and least negative datasets respectively. This contrasts with the sexual violence datasets which vary between 1% negative for vLNSH500, and 12% for vLTMT500. Twitter datasets are always more negative than those from Nexis, with vLTPG500 having almost twice as many concordances with negative prosody as vLTMT500, the next most negative.

The answer to RQ2B is that six themes directly mapped to social actor codes: Activation, Appraisal, Beneficialisation, Indifferentiation, Subjection, and Suppression. Within these themes my analysis found the subthemes: Abuse, Causes, Consequences, Definitions, Denial, and Speaking out. The analysis showed that those affected by the pay gap are seldom referred to directly, and those advocating for closure of the pay gap are often met with aggression. The themes within positive representations (section 5.4) are predominantly from Nexis and they portray the pay gap as a serious problem that needs to be corrected. In contrast, the negative representations (section 5.5) are often from Twitter, and are characterised by either: denials of the pay gap's existence; or references to causes and consequences that are analogous to nature versus nurture discussions. Throughout there is confusion between equal pay and the gender pay gap.

The answer to RQ2C is that the themes of pay gap representations additionally include three themes absent from the sexual violence datasets: Abuse, Definitions, and Denial. Those disadvantaged by the pay gap are held accountable for making the wrong "choices" and for personal failings in the workplace. The sexual violence and pay gap datasets have in common the promotion of personal responsibility and "women's choices", which at times manifests as victim-blaming, suggesting shared individualistic and neoliberal viewpoints that do not acknowledge social barriers.

### 8.1.3 RQ3

Research question 3, analysed in Chapter 6, stays with the pay gap but uses the full corpora as the data sources.

RQ3, How are truthfulness and legitimacy attributed in the pay gap discourses, and how is this different from that seen in the sexual violence datasets? (Chapter 6)

The answer to RQ3 is that within the pay gap corpora it is the existence of a gap or its validity as an injustice that is argued, while within the sexual violence corpora the victims/survivors of sexual violence are personally accused of lying, deceit and defamation (Chapter 6). This derives from the pay gap being a statistical phenomenon, and pay gap deniers commonly complain that its statistics are based on faulty data. Confusion arises from misunderstandings of equal pay and the pay gap. In contrast, as a crime against the person, crimes of sexual violence are weakened and deflected by discrediting the accuser. In both cases Just World Theory (JWT) explains that by attaching fault to other people's personal behaviour, third parties and society protect their perception of the world as generally a safe place, where people get what they deserve. JWT also influences who receives sympathy and support, providing a mechanism for policing dominant cultural values and maintaining the status quo.

### 8.1.4 RQ4

The final question, RQ 4, refers to both sexual violence and pay gap discourse in addressing the issues of labelling as a VICTIM or SURVIVOR, and what that means for those who are muted, silenced or missing from the discourse (Chapter 7).

RQ4 How does delegitimation manifest for VICTIMS and SURVIVORS? This question is addressed in two parts.

RQ4A, Is the SURVIVOR label more beneficial than the VICTIM label for individuals, and what are the implications for society? (sections 7.3 & 7.4)

RQ4B, Who are the missing? (sections 7.5, & 7.6)

RQ4A addresses a complex situation. A broad-brush answer is that SURVIVOR may be more advantageous to the individual in the short term. However, the general trend towards SURVIVOR away from VICTIM risks weakening support for, and delaying healing of the victimised, and obscuring systemic problems and perpetrators.

If an individual has healed and truly put their victimisation behind them, then a SURVIVOR's perceived strength and agency may be appropriate and empowering, and likely to attract cultural approval. However, if the label is taken up before healing is complete, having to live up to the expectation of strength and agency while still carrying trauma may become a heavy burden, suppressing trauma and delaying recovery (section 7.4.1.1). Additionally, if the victimised needs to take legal action then being labelled a VICTIM is a pre-requisite to having a perpetrator within the legal system (section 2.5.2).

The toxification of VICTIM, and valorisation of SURVIVOR since the 1970s has happened in parallel with a shift towards neoliberalism in many developed countries (section 7.3.4). Neoliberal citizens are viewed as responsible for their personal success and safety, which aligns with the characterisation of the strong and agentic survivor. Hence, an ideal neoliberal

citizen resolves their own issues, does not ask for help, and restores economic normality as quickly as possible (section 2.5.6). This view is blind to social inequities and barriers, discriminating against the least privileged in society. Moreover, with every person who positions VICTIM as being a lesser choice, each public statement rejecting VICTIM in favour of SURVIVOR further tarnishes and weakens VICTIM for those who need to challenge injustice.

RQ4B is something of a paradox because it is not possible to prove a negative. However, there are individuals and groups who could reasonably be expected to be present within the discourse, but aren't. A creative approach was thus taken using a process of elimination, starting with the exclusion of the visible demographics. Beyond linguistic exclusion, the task is compounded by the high numbers of individuals likely to be missing due to under-reporting of sexual violence, and lack of awareness of the pay gap.

Within the sexual violence corpora the missing are those who do not report, do not speak out informally, and do not complete anonymous surveys (sections 2.5.7 and 7.6). Others are unaware that their victimisation is a crime, or fear retaliation, especially in cultures where sexual violence is normalised (section 2.5.9). Those who have spoken out, may be delegitimised, or judged as undeserving in some way, resulting in their claim being taken less seriously or disbelieved. Effectively this silences their voices, and also deters others from speaking out. The obscuring and muting of victims allows perpetrators to continue with their crimes, and supports the JWT view that such crimes are rare and the world is a safe place (section 2.4.4).

The missing from the gender pay gap discourse may be unaware of their disadvantage due to the widespread confusion about the difference between the pay gap and equal pay (section 2.2.2). Despite long established legislation, failures of equal pay still feature in the media. However, the definition of the pay gap is vastly wider than equal pay, including issues such as: career segregation, glass ceilings, incorrect and limiting gendered stereotypes, parental leave, unpaid work, and gendering influences on children (section 2.2.2). These issues are complex and, without an obvious perpetrator, leaves many unaware that they have been disadvantaged (section 2.5.1.1). After accounting for the demographics that were found within the corpora and other sources referenced in this study, it suggests that the missing are likely to be workers in non-unionised, traditionally female roles, working part-time and/or zero-hours contracts, in sectors lacking pay transparency, and with low personal expectations of justice (section 7.5). However, that is not to say that employers or employees are conscious of bias. Rewards Expectation Theory suggests that both women and men are biased in evaluating others worth according to their pay and status, the highest ranking and best paid being viewed as most capable, maintaining the view that men's work as more valuable (section 2.4.5).

## **8.2 Reflections**

I started the present study with the general intention to try to find out what was happening in discussions about the pay gap and sexual violence, and whether there was any connection between the two. Reaching adulthood just behind the crest of second wave feminism, I was aware that both were feminist issues frequently reported in mainstream media, and the lived experience was discussed as an everyday issue by female friends and relatives. At this point, I hadn't known how the two issues might be connected. Individuals represented

within the sexual violence and pay gap corpora shared several van Leeuwen social actors, and they provided a useful way to categorise and discuss patterns, although often using them in opposite ways, and these social actor types provided a way into the data.

First indications of a link between the issues came from journalistic opinion pieces (section 2.2.3) discussing workplace discrimination, which suggested that sexual harassment could hinder women's career progress, while lack of progress into senior roles could prevent sexual harassment being addressed. Women's lack of power in the workplace allows these mechanisms to function, and undermines their attempts to change the situation.

Governmental reporting was also starting to identify a relationship between sexual violence and the pay gap. Dawn Butler (2020) speaking on International Women's Day, identified "intractable structural inequalities" as underpinning both sexual violence and pay disparity, fifty years after preventative legislation was enacted in the UK.

After struggling for the first half of my timeline to find representations within the pay gap corpora, I realised that this was exactly the issue: many of the disadvantaged people were missing and therefore silenced (section 7.5). I also identified various forms of silencing associated with the sexual violence texts although here taking different forms (section 7.6). The concepts of power and silencing, post-feminism and their relationship to neoliberal politics, crystallised to provide strong possible explanatory mechanisms (sections 7.3.4 and 7.3.6).

The methodology used in the present study adds weight to the understanding that real world data is messy, and Twitter's data is especially messy. Social media content is written and read on the fly, generally without a second reading or edit, and a single conversational turn may be split between multiple postings. However, the writer's meaning is only half of the equation, with the reader's interpretation making up the other, and this is influenced by the platform's algorithms (section 2.5.3.1). I have taken the view that a qualitative analysis is the best way to identify a meaning, while remaining cognizant of FPDA's caution against fixing on a definitive meaning. Within social media, algorithms match and strengthen users' established preferences, so users are more likely to be shown postings in line with their existing beliefs, fostering polarisation. The Twitter downloads for this study were taken from a new account with no history, so a variety of viewpoints were gathered. Moreover, in the real world tweets speaking for or against victims/survivors of sexual harassment or the pay gap, are more likely to be read by users with a Twitter history of similar tweets. Furthermore, Twitter's algorithms are biased towards negativity (section 3.2.2) and each user feed may contain different tweets, so the idea of a single audience or single meaning is confounded.

With 10% of writers responsible for 80% of tweets, (section 3.3.1) the extreme right-leaning and negative postings are likely to involve a small percentage of Twitter @useraccounts. However, they may be very visible and highly proliferated to influence many people (section 3.2.1.1). Thus a user on Twitter is likely to see higher frequencies of such posts, than found in these deduplicated corpora, which prioritise the range of voices.

The social context is thus positioned as political, and operating unconsciously through Twitter's provision of a safe space to speak out but which also enables retaliation, protecting the hegemony and the status quo (section 2.5.3.1). For example, the single initiating tweet of the #MeToo hashtag, immediately went viral and eventually resulted in a sequence of high

profile court cases, but it also generated a negative backlash. Even when writers lack authority, a stream of constant negative utterances working together at a low level can undermine and intimidate those wishing to speak out (section 6.1) effectively policing normative values and behaviours.

### 8.3 Limitations

There is an inescapable element of subjectivity running through this study, from the selection of topic, through to choice and application of theories. The influence of the researcher's interests, intuition, and prior knowledge is inevitable (section 2.3.1) and I outline my characteristics most likely to have created bias (section 3.6). This study positions itself as an interpretation of this data, while at the same time acknowledging that there are other interpretations. Other researchers may have coded differently (sections 4.1 and 4.3.1), and taken different views of themes, for example Cause and Consequence (section 5.4.1.2). Additionally, the choice of a statistical measure other than LogDice (section 3.3.8.2) would have resulted in an entirely different study. While all of this could appear to diminish the value of this study as representing but one voice, it also underlines the importance of every single voice, none of which are more right than any other. Additionally, this study tries to be self-reflexive and uses frequency, and procedures that interpret those frequencies as a series of constraints on subjectivity. While interpretations may be subjective, the frequencies observed are not.

This study represents my exploration of a subject and data that I specifically identified as being of personal interest. When I started to find a relationship between sexual violence and the pay gap I narrowed in on that, consciously setting aside some features as beyond my scope. Other features may have been overlooked entirely. Further potentially insightful texts may have been lost in data cleaning, random selections, and pruning corpora down to size.

From the outset the intention of the study was to capture the broadest range of voices possible. This led me to exclude re-tweets during data harvesting, and remove the large numbers of near duplicates. Thus frequencies and statistical patterns appearing in these corpora will not be representative of the features displayed within Twitter. Specifically, this biased the data towards more positive content. In contrast, the more common victim-blaming accounts and denigrating tweets have more followers and are more frequently retweeted than those supporting the victimised (section 2.5.4). A visitor to Twitter in 2019-2020 would have encountered a much higher percentage of tweets challenging and denigrating the victimised, but they may not have noticed that many originated from the same source.

This study was primarily interested in what interlocutors do with language and the effect it may have on others. Thus, coding of +/- prosody was based on a subjective assessment of whether upon reading a concordance a victimised person was likely to feel encouraged to speak out (positive) or feel intimidated or discouraged and less likely to speak out (negative) (section 4.2). Sometimes victim-denigrating quotes are called out in tweets, and these were coded as supportive of victims and therefore positive. A narrower reading including the quote in isolation would have coded these as negative prosody. Veracity was coded according to whether or not the subject was represented as being truthful. This was sometimes achieved by implication, for example by stating that others were the *real* victims,

the subject of the sentence by implication was not a *real* victim and therefore bogus. A strictly explicit interpretation would have resulted in another coding. Coding of van Leeuwen social actor types was sometimes ambiguous and complicated, with multiple and sometimes contradictory types co-existing (section 4.3). In the context of tweets that are written and read in seconds, I decided that coding according to an early impression was closest to a non-academic reading experience on Twitter (section 4.3.2.1). Adopting a wider reading my study extended beyond concordances to their original texts, linked articles, and extra-textual cultural knowledge. Other researchers may have taken a different approach and/or made different decisions. For example, a narrower reading, excluding resources beyond the concordance, or with different cultural understandings, may have resulted in other codes.

Six 500 concordance datasets were manually coded, within which 29 social actor types were identified. In order to devote more analysis time to the most frequent social actors, a threshold of 100 occurrences across the vL500 datasets for either VICTIM, SURVIVOR, or PAY GAP was chosen (section 4.3.1). A different threshold for analysis may have produced different results.

The missing were defined as those who I would have expected to be present in the texts, but weren't. Other researchers with different expectations may have identified different groups as missing (section 3.2.5). Alternatively if a researcher believed there is no pay gap, no one could be disadvantaged by it, and there would be no debate from which to be absent.

#### **8.4 Further research directions**

Several potential topics came to light while conducting this study that I found interesting and may have research potential, but were beyond the current scope. For example, future research could usefully explore the victimisation of specific demographics such as ethnic groups, or within particular environments, defined for example by workplace, sport, or politics. As VICTIM and SURVIVOR are gendered terms, discourses of victimised men would make a particularly interesting topic.

The decision to prioritise the range of voices resulted in heavy deduplication and a skewing of data towards victim-supportive tweets. Another study could omit deduplication to foreground the true size of peaks and troughs of statistical features that users would have experienced on Twitter at the time. This would likely result in fewer features, but more dramatic frequencies for those remaining. A study of parallel deduplicated and unduplicated corpora would also make an interesting comparison.

Another interesting issue that emerged was the interpretation of withdrawn legal cases and "not guilty" verdicts. Within this study both were often stated as proof of deliberate false claims, and accusers in some cultures having been counter-sued for defamation when their case did not end in a guilty verdict. Exploring discourse around terminology such as *guilty*, *not guilty*, *acquitted*, and *insufficient evidence* could be enlightening.

Pilot WordSketches for this study resulted in far more references to US politicians than expected, and although incidental to this research, it would be an interesting topic for another study to centre on accusations of personal wrongdoing used in politics.

## 8.5 Contribution and implications

This dissertation builds on existing corpus linguistic research into gendered discrimination by focusing on both sexual violence and the less commonly studied gender pay gap. The source texts were gathered from 2019 to 2021 when both topics were receiving much public attention following on from the #MeToo<sup>8</sup> movement and the BBC pay scandal<sup>2</sup>, suggesting a high level of public interest in the topics. The contribution of this study has four novel elements, the:

- use of large-scale deduplicated corpora that enable the capture of multiple voices within sexual violence and pay gap debates, as advised by Feminist Post-structural Discourse Analysis (FPDA);
- application of Just World Theory (JWT) to sexual violence text to help explain the resilience of victim-blaming;
- identification of a hermeneutical injustice within pay gap corpora;
- analysis of corpora to strengthen Feminist Legal Theory (FLT) claims with quantitative evidence from both sexual violence and pay gap corpora.

These four elements are now discussed.

### 8.5.1 Prioritising the range of voices

This study has used Feminist Post-structural Discourse Analysis (FPDA) concepts and applied them to the written texts of these corpora (section 2.3.2). FPDA was designed for the analysis of spoken language, and promotes a broad CDA type approach that embraces the multiple and contradictory voices within a discourse, and acknowledges shifting power relations.

This study maximises the “multiple voices” by use of deduplication making its six 2.4 million word corpora qualitatively different from un-deduplicated corpora that may be dominated by a few writers. This is because within Twitter 80% of original tweets, at the time of data harvesting, were produced by only 10% of user accounts (section 3.3.1). When visiting Twitter, or downloading in bulk via the API, users are likely to predominantly find retweets, and tweets containing text copied from earlier tweets. Which texts are promoted in user feeds, to drive retweeting and copying, is shaped by the platform’s algorithms that favour negative and extreme posts (section 3.3.2). Thus the algorithms increase the number of tweets showing negative prosody that are delivered to the user screen or API, creating an online world that exaggerates negativity. By creating corpora which exclude retweets and near duplicates, and limit posts from single user accounts, a wider range of the “multiple voices” is captured. At the same time, this approach minimises the effect of any bot generated texts.

Where corpus linguistics research is carried out for gendered discrimination topics, deduplication is generally discussed in relation to news articles or the exclusion of retweets, rather than identifying and removal of closely similar text. I believe it is the first study to deduplicate this extensively to embody the FPDA stance of embracing multiple and contradictory voices in sexual violence and pay gap corpora.

Those responsible for the design and the safety functions of Twitter should care about this because it is within their power to reduce negative representations. However, Twitter's business model relies on advertising, and users engage more with negative and extreme content, providing motivation for Twitter to encourage negative postings. Consequently, governments need to provide regulatory safeguards to support change in a more positive and socially responsible direction.

### **8.5.2 Survivor and victim, RQ1**

The present study contributes to our understanding of how the people who are victimised by sexual violence are represented by using quantitative data. Within its own datasets this study found that *victim* is represented as negative by three times more voices than *survivor*, and additionally it has shown how *victim* and *survivor* representations differ. There is a strong body of research on the use of the terms *victim* and *survivor* from, for example criminology (section 2.5.1), legal feminism (section 2.5.2), or focused on news articles (section 2.5.3). However, this is either qualitative, lab-based, or includes social media data that may be dominated by a few voices and their message.

The wider research has identified negative associations for both the *victim* and *survivor* labels, and this study extends understanding of these into large new corpora and the datasets based on them. This study has found negative and positive connotations of both *victim* and *survivor* labels (sections 7.3 and 7.4) and additionally supports and expands understanding of the issues by providing quantitative evidence. Existing research reports biased use of terms such as *alleged* (section 7.3.5), *real*, and *true* (section 3.4.2.1 and 3.4.2.2) and these terms are also found in this study creating negative prosody around *victim*. This study has specifically analysed *victim* alongside *survivor* to be able to both support earlier research in finding negativity associated with *victim*, but also to extend this to confirm that *survivor* is not associated with negativity in this way. This study's contributions will now be discussed in more detail.

In this study analysis has shown strong themes of victim-blaming both directly counter-accusing those speaking out with statements such as *the woman is playing the victim just to get money* (Example 56), and indirectly where such statements are called out. Because *victim* is associated with weakness and passivity, it is interesting that victim-blaming is a very strong theme for victimised Activation social actors. Activation was the next most common social actor type after Beneficialisation and Subjection. However, although agency is usually a desirable attribute for an individual, here the agency is positioned negatively as the cause of their victimisation. I have used Just World Theory (JWT) (section 2.4.4) to propose that victim-blaming may be used by a third party to protect their own belief in a fair world. JWT also operates at societal level, seeking to protect the general perception of safety within the social world. Exposure to victim-blaming discourse may cause victims to internalise the discourse and self-blame, feeding into the under-reporting of crime (sections 2.5.2, and 2.5.9). Moreover, *victim* implies a perpetrator, so holding the *victim* culpable removes responsibility from both the perpetrator, and the systemic structures that allow discrimination to happen.

The inverse is the more frequent association of *survivor* with positive prosody, as has been suggested by journalists (section 2.5.6) speculating that *survivor* is socially preferred over

*victim*. This is not a straightforward choice between two synonyms, because, in addition to having legal meaning, *victim* carries the powerful prosody of wrongs carried forward from centuries of victimisation, which *survivor* lacks. While those adopting *survivor* gain a strong and agentic identity, they may be unaware of its risks. *Survivor* comes with neoliberal expectations of individuals owning their personal safety and recovery from any negative outcome, and also the obligation to speak out using a particular type of narrative (sections 2.5.6 and 7.3.4). This study found evidence of an association between *survivor* and the expectation to speak out, and to advocate for others.

This means that groups supporting and advocating for the victimised need to be wary of the potential down side of a positivity culture that encourages adoption of *survivor*. Inappropriate use of *survivor* may result in individuals under-reporting, minimising the harm, create a barrier to accessing support, and suggest an obligation to speak out and provide a positive narrative. In promoting *survivor*, *victim* becomes devalued as the “lesser” option, eroding its power to challenge injustice, and distancing both perpetrator and crime. Perhaps most importantly organisations supporting the victimised need to be aware that the long term cumulative consequence of a shift towards *survivor* may constitute enfranchised semantic drift (section 7.3.6) bolstering existing power dynamics and reinforcing discrimination. Ultimately, in a world where *victim* is entirely robbed of its sense of victimisation, discussing or challenging the associated crimes would be made more difficult. Acts of victimisation and any patterns of abuse would be less easily recognised by society, perpetrators, and victims themselves. Unrecognised and unacknowledged as harmed, any consequent difficulties or discomfort would become personal issues for the victimised to resolve themselves. This position maps closely to the idealised neoliberal citizen for whom victimhood is viewed as a personal failure, rather than an injustice to be addressed by society.

### **8.5.3 Those wronged by pay gap RQ2**

This study found no previous research that either identified a term for those disadvantaged by the pay gap analogous to *victim* or *survivor*, or pointed out the absence of such a word. Those disadvantaged by the pay gap, like those in the sexual harassment corpora are portrayed negatively, but in different ways. Most often it is the pay gap itself that is discredited, which indirectly discredits those advocating for its closure. When the disadvantaged are referred to in existing research, an adjective or short phrase is used to describe having been wronged, for example women disadvantaged by the pay gap may be positioned as “low-paid” (Amadjarif et al., 2020, p. 9), “economically vulnerable”, or “paid less than men” (Gould et al., 2016, pp. 18, 26). Here the disadvantaged person is described rather than given a universally understood label that enables individuals to recognise each other, and which would support group action (section 2.5.9). This is analogous to individuals presenting with a health condition before it is recognised as a diagnosis. Hermeneutical injustices have historically been associated with powerlessness, and in this situation women’s attempts to communicate their experience are hindered, risking miscomprehension. To my knowledge this study is the first to identify hermeneutical injustice in relation to those disadvantaged by the gender pay gap. This study employed a corpus linguistics and CADS analysis involving extratextual exploration, to search for and then confirm that no single word exists within the corpora of this study that represents the

disadvantaged as have been wronged. I have interpreted this as a hermeneutical injustice (section 2.5.9 and 5.7). Where the disadvantaged are described in neutral and generic terms, such as *women*, it could suggest that to be a woman is to be disadvantaged by the pay gap. This would align with the UK “Gender Equality Roadmap” (Government Equalities Office, 2019) which identifies multiple systemic biases and barriers that are encountered in normal female life events, and which combine to drive the gender pay gap.

This is important because a hermeneutical injustice suggests a deep rooted ideological problem. While the pay gap is identified by reporting that shows companies with poor pay gaps, the reporting positions the onus entirely on companies, for example, to address a glass ceiling. Reporting seldom suggests the structural bias and barriers that constrain the choices of girls and women before they reach that point. With no term to indicate the disadvantage, arguments of *women* being harmed are undermined, easily presented as vexatious or frivolous claims by a group that already has equality legislation.

This study also found evidence that some writers express neo-liberal and post-feminist views that the pay gap is either *fought for and won* (Example 154) or existing but justified *The pay gap is fair, not patriarchal* (Example 218). Others wrote that the pay gap is *political correctness gone too far* (Example 110). Whether or not people believe the pay gap needs to be closed, they may still be disadvantaged by it (section 2.2.2). Women internalising such views, who are unaware of their disadvantage, may conclude that any difficulties are personal problems. Being unable to identify a structural cause for the disadvantage, they are hindered in challenging it or speaking out, and thus may remain silent.

The corpora used in this study also provide evidence that the definition of the gender pay gap and its wider societal roots are poorly understood, even by mainstream media who regularly confuse the pay gap with equal pay. Pay gap issues, which include equal pay but also career segregation, the glass ceiling, and the motherhood penalty, are positioned as rightful and natural consequences of “women’s choices”, and are present as a strong theme in my data (sections 2.3.2, 5.5, 6.4). This suggests widely held ideological beliefs that underpin the construction of gendered identities for women and men and what they should and shouldn’t do, supporting the neo-liberal view of women as primarily home-makers and un-paid carers.

This study highlights that the discussion of the pay gap is seriously impeded by a lack of clarity over the terms *pay gap* and *equal pay*. Clarifying these terms should be a first step to addressing the pay gap. It should be undertaken by both governments and advocacy groups. The UK government should care about this because, firstly it committed to ending the gender pay gap in the “Gender Equality Roadmap” (Government Equalities Office, 2019). Secondly, it is within the government’s power to address some of the underlying causes of the pay gap identified in the Roadmap, for example by: reshaping of the school curriculum to prevent children absorbing gendered limiting beliefs; improving child care provision enabling mothers easier access to work; and introducing company reporting on the gap between the parental leave taken by men and women. In parallel, the government also needs to promote research into masculinities to ensure that boys develop a positive sense of identity and are not threatened by a gender equal society.

#### **8.5.4 Pay gap and sexual violence RQ3, RQ4**

While there is much research into sexual violence discourses, to the best of my knowledge this is the first corpus study to explicitly analyse and compare sexual violence with pay gap discourses. Cooke's (2024, p. 354) analysis of gendered inequalities in the Deaton Review found that women's economic and political power was "precursor and consequent" to reducing domestic violence and sexual harassment. This clearly indicates that a relationship between the two exists and that it is important.

The present study uses a combination of CADs and corpus methods to add quantitative evidence to the more qualitative FLT theories by examining both sexual violence and pay gap discourses. The study's large-scale corpora contain a range of victim-blaming strategies that give insights into discourses beyond the court room that are unconstrained, unedited, and spontaneous responses to the topics.

FLT interprets the use of rape myths to victim-blame as arising from the legal system's position within a gendered and discriminatory institutional context. As a crime, sexual violence is defined in the wording of laws and tested by the formal legal process, and FLT has produced a significant body of research into language in and of the legal system (section 2.4.3). However the use of language within the court room constrains victims' testimony (section 2.4.3), and conviction rates are poor (section 2.5.2). In this context many victims have embraced the opportunity to share their un-edited stories on social media (section 2.5.3.1) and my study now extends research into sexual violence discourses posted on Twitter. FLT is unable to include pay gap discourses because the gender pay gap is not illegal, however this study is able to compare the sexual violence discourses with those of the gender pay gap using its quantitative and qualitative approach.

To the best of my knowledge this is the first study to visualise FLT theories from a JWT perspective. JWT positions society and individuals as unconsciously using victim-blaming to protect their belief in the world as a safe place, and their own safety within it if they do the "right" things and avoid the "wrong" things. The psychological need to feel safe provides an important explanation for the resilient nature of rape myths despite decades of evidence disproving them and the justice system's guidance on avoiding them.

Many of the themes found in sexual violence corpora also exist for the pay gap. However, they tend to manifest in different ways. For example, victim-blaming within the pay gap corpora takes a different form, and is usually targeted at generalised *women* rather than individuals. Veracity was found to be a strong theme with the existence and evaluation of the pay gap featuring heavily. Here justification is achieved by blaming women's choices, lack of effort, and inferior workplace skills (section 6.7). This contrasts with the victim-blaming in the sexual violence corpora that directly implicates the victim, for example in workplace cases victims may be accused of having ulterior motives using the quid pro quo myth (section 3.2.4.2). Victim-blaming by either strategy attributes agency to the social actor that may not exist in reality, especially when gendered social barriers and power relations restrict options.

By blaming individuals for their own victimisation or denying that harm has been done, there is no claim against a perpetrator and no injustice to be addressed, leaving the victimised individual isolated and responsible for their own recovery using only their own resources.

Thus the individual may be effectively silenced, and others witnessing these discourses may be deterred from speaking out. This mechanism aligns with a neoliberal ideology that foregrounds personal responsibility for individual safety and success, while ignoring any structural barriers encountered by particular groups.

The Crown Prosecution Service (n.d.) should care about the resilience of victim-blaming, because their guidelines explain that rape myths are false, suggesting the CPS has committed to eradicating them. However, JWT explains how society and individuals may rely on rape myths to feel safe making such beliefs resilient, so that the CPS also needs to monitor compliance. Additionally, the CPS is in a position take steps to provide mandatory rape myth training for legal professionals and journalists. Rape myths will not easily be removed from everyday conversation while they are legitimated in the court room.

As a broad societal problem the government needs to intervene at a high level. The multiple roots of the pay gap are complex and woven into society and their influence is cumulative over time. Consequently, any proposed changes to address the issues identified in the Gender Equality Roadmap would require specific analysis to ensure that the interventions are likely to deliver the necessary results. Prior to that however, clarification is needed for the terms *pay gap* and *equal pay* discussed earlier.

The contributions of this study matter on a number of levels. Firstly, discrimination based on protected characteristics is illegal and must be challenged. Secondly, there are damaging personal consequences for individuals prevented from achieving their full career potential, or living with the aftermath of sexual violence. Thirdly, negative and extreme texts aligning with neoliberal ideology are being systematically amplified by Twitter in a way that compounds the harm.

## **8.6 Concluding remarks**

It is easy to lose sight of the fact that the majority of tweets within the corpora are positive. Within the most negative dataset, vLTPG500, 67% of concordances are positive. Moreover, most of the tweets come from the top 10% of tweeters (section 3.3.1) so the extreme negative prosody seen in this study is likely to come from a relatively small number of individuals. However, because Twitter's algorithms favour negative and extreme postings, and tailors individual feeds according to previous searches (sections 3.2.1.1 and 3.3.2) individual messages can go viral giving them huge impact. Thus a small number of writers can flood Twitter with negative representations.

The analyses of this study, combined with the research described in Chapter 2 create a sense of a powerful ideology, with most of the mechanism sitting below the level of awareness, but functioning to police the ideology's values. The term SURVIVOR appeals to an individualistic right-leaning ideology because it foregrounds the personal characteristics bringing about the survival. In contrast, VICTIM implies no such personal qualities because the individual is entirely on the receiving end of the victimisation (section 2.5.3 and 7.3.4). A victim then is not responsible for their own victimisation, and positioned as lacking agency should be deserving of support and justice, but this is out of alignment with an individualistic culture. Where systematic biases harm non-dominant demographics, this constitutes discrimination. However, discrimination suggests a political rather than a personal issue, which may pose a threat to an individualistic culture. Language and culture have a reciprocal

relationship (section 2.4.2) with individual language choices aggregating into general shifts in semantic prosody, which then have broader consequences for society. For example, adopting a SURVIVOR label in preference to VICTIM may represent an ideological shift that is likely to maintain the status quo, or even provide benefit to the dominant group, rather than helping victims challenge injustice (sections 7.3.3 and 2.5.10). A weakened VICTIM label diminishes the vocabulary available for discussing sexual violence. For instance, the use of SURVIVOR to challenge injustice because we cannot talk about having been *survivorised* by *survivoriser*, distancing the crime and the perpetrator (section 7.3.4). Furthermore, VICTIM itself is a powerful word with strong semantic prosody that associates it with wrong-doing accumulated over centuries (sections 2.5.10 and 7.3.6). While SURVIVOR portrays agency and courage, it loses sight of the original wrong-doing.

In common with those disadvantaged by sexual violence those within pay gap concordances are living in an individualistic society that views each worker as responsible for their own success and any difficulties are a personal issue. Their engagement in the workplace might be inhibited by structural issues such as lack of childcare, parental leave, or unpaid care work. The situation is complicated by a significant amount of the discourse being at cross purposes which may account for much of the emotion demonstrated. While the communications of most international bodies, governments, and large companies show a correct understanding of the terms there is much other written material circulating that uses the terms *pay gap* and *equal pay* either incorrectly or even interchangeably. This study found evidence of *equal pay* terminology being used to claim that there is no *pay gap*. Such denials render invisible gendered workplace barriers such as: glass ceilings, career segregation, and child care issues (section 5.5). Within the corpora those speaking out on the pay gap are represented with negative prosody and as: lying or deluded (section 6.4); unacceptable (section 6.5); or viewed with contempt (section 6.6). All of which either eliminates or mutes any claim of victimisation. By ignoring structural barriers disadvantaged individuals are rendered invisible because their disadvantage is explained as the natural result of “women’s choices” (section 6.7). Where discrimination is denied there is no victimisation, and those claiming victimisation become invisible.

There is evidence suggesting that a large number of individuals are missing from both sexual violence and pay gap debates (sections 7.5 and 7.6). These two groups of missing individuals overlap and have issues in common, but they also show differences. Within the pay gap debate, the absence of those we could have reasonably expected to be there, is demonstrated most clearly by the absence of a word to describe the wronged. Many of those missing from the pay gap debate are likely to be unaware themselves, with their disadvantage originating far beyond the workplace. In contrast, most of those missing from the sexual violence debate are very aware of their victimisation, but may be silenced by fear of disbelief, retaliation, shame, or the inability to emotionally process the abuse.

This study has shown how language with negative prosody is used in discussing both the pay gap and sexual violence harms and delegitimises the disadvantaged individuals, or removes them from the debate entirely. Both issues are underpinned by deeply embedded power relations, and a cultural worldview that holds individuals responsible for their own misfortunes, and for dealing with the consequences.



## Appendix A – Rape myths

Extract of “Table outlining a range of myths, their implications and how they may be addressed” (Crown Prosecution Service, n.d.).

Myth	Implications
1. Rape is always violent or involves physical force. (FALSE)	<ul style="list-style-type: none"> <li>• Ignores the reality of rape.</li> <li>• Disbelieves and invalidates the experience of the victim.</li> <li>• Disregards elements of power, control and humiliation in rape.</li> </ul>
2. Rape most commonly occurs between strangers in dark alleys. (FALSE)	<ul style="list-style-type: none"> <li>• Assumes that home is always a safe place and that rape cannot happen between parties that know each other.</li> <li>• Implies that rape can be prevented by avoiding certain places and therefore blames the victim.</li> </ul>
3. Prostitutes / sex workers cannot be raped. (FALSE)	<ul style="list-style-type: none"> <li>• Provides an excuse for sexual abuse.</li> <li>• Disempowers prostitutes / sex workers.</li> <li>• Assumes that prostitutes / sex workers consent to any and all sex under any circumstances, with anyone.</li> </ul>
4. You cannot be raped by your husband or partner. (FALSE) 5. The victim had previously consented to sex with the accused a number of times so s/he must have consented. (FALSE)	<ul style="list-style-type: none"> <li>• Disbelieves the victim and dismisses their experience;</li> <li>• Supports an ideology of male entitlement to sex and female subordination within intimate relationships and more widely.</li> </ul>
6. If your culture condones, or is perceived to condone, marital rape, underage “sex”, or forced marriage, then you should not be upset about it/it does not matter as much/it’s more of a grey area. (FALSE)	<ul style="list-style-type: none"> <li>• Condones rape due to cultural reasons.</li> <li>• Creates further barriers to accessing support and justice.</li> </ul>
7. Sexual abuse at the hands of a perpetrator which took place when the victim was a child has no bearing on the issue of consent if the same parties go onto engage in sexual activity as adults. (FALSE)	<ul style="list-style-type: none"> <li>• Context, background and history are irrelevant factors when considering consent.</li> <li>• Disregards elements of power, aggression, violence, control, entitlement and humiliation in rape.</li> </ul>
8. Rape is only a crime of passion. (FALSE)	<ul style="list-style-type: none"> <li>• Romanticises rape and implies that rape is the same as sex.</li> <li>• Assumes that rape is impulsive, unplanned and about uncontrollable ‘passion’.</li> <li>• Assumes men to be incapable of delaying gratification or controlling sexual urges.</li> </ul>
9. When it comes to sex, men have a point of no return. (FALSE)	<ul style="list-style-type: none"> <li>• Disregards elements of power, aggression, violence, control and/or humiliation in rape.</li> <li>• Deprives a victim’s right to choose.</li> <li>• Attempts to remove the responsibility for the rape from the rapist.</li> </ul>

## Sexual violence and the gender pay gap

Myth	Implications
10. Young adult men should not be convicted as they have their whole lives ahead of them / have good character references. (FALSE)	<ul style="list-style-type: none"> <li>• Assumes that law does not apply equally to all.</li> <li>• Implies that the implications for the offender are more important than the impact of rape on the victim as well as wider justice and safety considerations.</li> <li>• Implies that young adult men are less able to understand consent.</li> </ul>
11. He was satisfying demands for BDSM / choking / aggressive sex. This cannot be not rape. (FALSE)	<ul style="list-style-type: none"> <li>• Assumes consenting to BDSM is consenting to any type of violence, abuse or rape.</li> </ul>
12. Only gay men rape other men / only gay men get raped. (FALSE)	<ul style="list-style-type: none"> <li>• Reinforces homophobic fears and prejudices.</li> <li>• Creates the illusion of safety for straight men.</li> <li>• Results in under-reporting of rapes on men.</li> </ul>
13. The victim provoked rape and implied consent simply by their dress / flirtatious behaviour. (FALSE)	<ul style="list-style-type: none"> <li>• Stigmatises and blames the victim.</li> <li>• Assumes that if someone draws attention to themselves, they are seeking sex, consenting to sex or 'deserve to be raped'.</li> <li>• Asserts that women cannot dress / behave how they want - that it's for others rather than their own preferences.</li> <li>• Attempts to remove responsibility from the rapist by taking away their agency - that they are so provoked by clothing or behaviour they cannot help themselves.</li> </ul>
14. If you send sexual images or messages prior to meeting someone, then having sex is inevitable. (FALSE)	<ul style="list-style-type: none"> <li>• Attempts to excuse rape and blame the victim.</li> <li>• Assumes that if someone sends a sexual image of themselves, they are seeking sex or 'deserve to be raped'.</li> <li>• Assumes sending of a message or messages automatically means physical sex is wanted any time, under any conditions.</li> <li>• Assumes that someone cannot change their mind, even if they had initially been interested in sexual activity.</li> </ul>
15. If you voluntarily attend someone's house after a date or night out, you obviously want sex and consented to it by going there. (FALSE)	<ul style="list-style-type: none"> <li>• Assumes that if a person goes to someone's house after a date or night out, they are looking for sex or deserve to be raped.</li> <li>• Assumes that rape is impulsive and unplanned.</li> <li>• Assumes that someone cannot change their mind, even if they had initially been interested in sexual activity.</li> </ul>
16. If you drink alcohol or use drugs then you have made yourself vulnerable to being raped and you bear the responsibility. (FALSE)	<ul style="list-style-type: none"> <li>• Disregards the law on consent, including capacity to consent and reasonable belief that someone was consenting.</li> <li>• Assumes that being under the influence of alcohol / drugs means that someone is looking for, or willing to have, sex.</li> <li>• Assumes that if someone is looking for sex, they therefore consent to anything that happens or are deserving of anything that happens.</li> </ul>

Myth	Implications
17. If you meet men online or through hook-up apps you are consenting to sex and should be ready to offer sex. (FALSE)	<ul style="list-style-type: none"> <li>Assumes that if someone goes on dating / hook-up apps or websites then they are looking for sex or deserve to be raped.</li> </ul>
18. If you have lots of sex, including with different people, then you are promiscuous and 'deserve what you get' and are not harmed by rape. (FALSE) 19. If someone has truly been raped then they would never seek, or want, sex soon afterwards. (FALSE)	<ul style="list-style-type: none"> <li>Assumes that if someone has multiple partners then they are looking for sex with anyone, deserve to be raped, or are less impacted by rape.</li> <li>Assumes that victims of rape behave in a certain way.</li> </ul>
20. You can tell if someone has 'really' been raped by how they act afterwards. (FALSE) 21. Real rape victims are always visibly distressed when describing what happened to them. (FALSE) 22. A real rape victim would never freeze when attacked, they would fight back. (FALSE) 23. A real victim of rape would never be able to carry on with their normal life – go to work, take children to school etc. 24. If the victim didn't scream, fight, or get injured then it could not have been rape. (FALSE)	<ul style="list-style-type: none"> <li>Disbelieves and re-traumatizes the victim.</li> <li>Invalidates the victim's experience and individuality.</li> <li>Assumes all victims behave in the same way/in a pre-determined way, during or following rape.</li> <li>Discourages the victim from seeking help.</li> </ul>
25. If the victim didn't complain to the police immediately it can't have been rape. (FALSE)	<ul style="list-style-type: none"> <li>Disbelieves and re-traumatizes the victim.</li> <li>Discourages him or her from seeking help.</li> <li>Assumes all victims behave in a uniform, pre-determined way.</li> <li>Ignores known trauma responses.</li> </ul>
26. If you don't say 'no', it's not rape. (FALSE)	<ul style="list-style-type: none"> <li>Disregards elements of power, aggression, violence, control and humiliation in rape.</li> <li>Goes against the legal definition of consent.</li> </ul>
27. Only young/ attractive people get raped. (FALSE) 28. Strong / independent / powerful / older people don't get raped. (FALSE)	<ul style="list-style-type: none"> <li>Assumes that only young or 'attractive' people are raped.</li> <li>Disregards elements of power, aggression, violence, control and humiliation in rape.</li> </ul>
29. The victim's race / religion / background is responsible for the rape. (FALSE)	<ul style="list-style-type: none"> <li>Condone rape due for racial / religious / cultural reasons.</li> <li>Creates further barriers to accessing support and justice.</li> <li>Can deploy racial and religious prejudices and stereotypes to blame the victim and disregard the facts of the case.</li> </ul>
30. A real victim would always be able to provide a clear and coherent account of being raped. (FALSE)	<ul style="list-style-type: none"> <li>Reinforces stereotypes of women as untruthful.</li> <li>Disregards the impact that trauma has on memory.</li> <li>Assumes that human beings have perfect memory recall.</li> <li>Discourages him or her from seeking help.</li> </ul>

Sexual violence and the gender pay gap

Myth	Implications
<p>31. Inconsistencies in accounts provided by a victim always means they lack credibility as a witness. (FALSE)</p> <p>32. Where a victim has consumed alcohol or drugs prior to an incident s/he will always be an unreliable witness as their evidence won't be accurate. (FALSE)</p>	<ul style="list-style-type: none"> <li>• Intoxicated victims are inherently unreliable witnesses.</li> <li>• Disbelieves and re-traumatizes the victim.</li> <li>• Discourages victims from seeking help.</li> </ul>
<p>33. False allegations are common and women<sup>1</sup> always cry rape when they regret having sex or want to seek revenge. (FALSE)</p> <p>34. Other complaints of rape which have not resulted in successful prosecution outcomes always mean the victim lacks all credibility as a witness. (FALSE)</p>	<ul style="list-style-type: none"> <li>• Reinforces stereotypes of the 'vindictive woman'.</li> <li>• Reinforces stereotypes of women as untruthful.</li> <li>• Re-victimises and stigmatises the victim.</li> <li>• Undermines her support for seeking justice / access to courts.</li> </ul>
<p>35. Previous withdrawals of complaints, or previous reluctance to co-operate with a prosecution, always means the victim lacks credibility as a witness. (FALSE)</p> <p>36. Where the victim has previous convictions s/he always lacks credibility as a witness as a result. (FALSE)</p> <p>37. The victim has previous convictions or had told untruths about other matters and so can never be relied upon to tell the truth about rape. (FALSE)</p>	<ul style="list-style-type: none"> <li>• Reinforces stereotypes of the 'vindictive woman'.</li> <li>• Reinforces stereotypes of women as untruthful.</li> <li>• Re-victimises and stigmatises the victim.</li> <li>• Undermines her support for seeking justice.</li> </ul>
<p>38. Where the victim has a learning disability or mental health condition s/he always lacks credibility as a witness. (FALSE)</p>	<ul style="list-style-type: none"> <li>• Disbelieves and re-traumatizes the victim.</li> <li>• Disregards elements of power, aggression, violence, control and humiliation in rape.</li> </ul>
<p>39. If someone displayed signs of sexual arousal during abuse, the only conclusion is that they wanted and/or enjoyed it (FALSE)</p>	<ul style="list-style-type: none"> <li>• Disbelieves the victim.</li> <li>• Disregards how bodies and brains can work.</li> </ul>
<p>40. Gay men who attend sex parties and/or take drugs are asking to be raped. (FALSE)</p>	<ul style="list-style-type: none"> <li>• Assumes gay men are promiscuous and are attracted to all men.</li> <li>• Assumes that sex is not romantically significant to gay men - it is all something they regularly engage in on a casual basis.</li> <li>• Assumes gay men do not have long-term loving relationships - for gay men it is all about sex.</li> <li>• Assumes gay men always have penetrative sex.</li> <li>• Assumes men cannot be victims of sexual offences.</li> <li>• Assumes men who have sex with men are gay.</li> <li>• Denies any incidence of coercive control ie if man was coerced to attend/participate.</li> </ul>

**Appendix B – Corpus and dataset naming****Corpora**

24 month data ~ 2.4 M words	12 month data ~ 1.2 M words	Monthly corpora ~100 K words
TwitterSH	TwitterSH_1Yr	TSH_01_2019_04Apr → TSH_12_2020_03Mar
	TwitterSH_2Yr	TSH_13_2020_04Apr → TSH_24_2021_03Mar
TwitterMT	TwitterMT_1Yr	TMT_01_2019_04Apr → TMT_12_2020_03Mar
	TwitterMT_2Yr	TMT_13_2020_04Apr → TMT_24_2021_03Mar
TwitterPG	TwitterPG_1Yr	TPG_01_2019_04Apr → TPG_12_2020_03Mar
	TwitterPG_2Yr	TPG_13_2020_04Apr → TPG_24_2021_03Mar
NexisSH	NexisSH_1Yr	NSH_01_2019_04Apr → NSH_12_2020_03Mar
	NexisSH_2Yr	NSH_13_2020_04Apr → NSH_24_2021_03Mar
NexisMT	NexisMT_1Yr	NMT_01_2019_04Apr → NMT_12_2020_03Mar
	NexisMT_2Yr	NMT_13_2020_04Apr → NMT_24_2021_03Mar
NexisPG	NexisPG_1Yr	NPG_01_2019_04Apr → NPG_12_2020_03Mar
	NexisPG_2Yr	NPG_13_2020_04Apr → NPG_24_2021_03Mar

*Table 76, The coded corpora**Key: T = Twitter, N= Nexis, SH = sexual harassment, MT = #MeToo, PG = pay gap**for monthly corpora the 1<sup>st</sup> pair of digits records sequence of data capture events 01 to 24**the 2<sup>nd</sup> pair of digits plus 3 characters (e.g. 04Apr) identifies the calendar month the data was published*

## Sexual violence and the gender pay gap

### Datasets

The vL500 datasets 500 concordance coded datasets	The pilot datasets Initial 100 coded concordances (only used in methodology)	Based on corpus
vLTSH500 VICTIM	vLTSHPilot VICTIM	TwitterSH
vLTMT500 VICTIM	vLTMTPIlot VICTIM	TwitterMT
vLNSH500 VICTIM	vLNSHPilot VICTIM	NexisSH
vLNMT500 VICTIM	vLNMTPIlot VICTIM	NexisMT
vLTSH500 SURVIVOR	vLTSHPilot SURVIVOR	TwitterSH
vLTMT500 SURVIVOR	vLTMTPIlot SURVIVOR	TwitterMT
vLNSH500 SURVIVOR	vLNSHPilot SURVIVOR	NexisSH
vLNMT500 SURVIVOR	vLNMTPIlot SURVIVOR	NexisMT
vLTPG500 PAY GAP	vLTPGPilot PAY GAP	TwitterPG
vLNPG500 PAY GAP	vLNPGPilot PAY GAP	NexisPG
	vLTSH_101-200 VICTIM	TwitterSH
	vLTMT_101-200 VICTIM	TwitterMT
	vLNSH_101-200 VICTIM	NexisSH
	vLNMT_101-200 VICTIM	NexisMT
	vLTSH_101-200 SURVIVOR	TwitterSH
	vLTMT_101-200 SURVIVOR	TwitterMT
	vLNSH_101-200 SURVIVOR	NexisSH
	vLNMT_101-200 SURVIVOR	NexisMT

*Table 77, The coded datasets*

*vL = van Leeuwen, T = Twitter, N = Nexis, SH = sexual harassment, MT = #MeToo, PG = pay gap*

**Appendix C – Deduplication and cleaning formulae**

[https://www.vistawide.com/german/top\\_100\\_german\\_words.htm](https://www.vistawide.com/german/top_100_german_words.htm)

<https://www.thoughtco.com/most-common-french-words-1372759>

Col	Heading	Purpose	Formula
A	user_screen_name	Twitter metadata	N/A
B	time	Twitter metadata	N/A
C	text	Tweet text	N/A
D	user_description	Twitter metadata	N/A
E	len	Calculates the no. of chars	=LEN(C2)
F*	Left	Copies the first 150 chars	=TRIM(LEFT(C2,150))
G	no. of @	Counts the no. of @	=LEN(C2)- LEN(SUBSTITUTE(C2,"@", ""))
H	no. of #	Counts the no. of #	=LEN(C2)- LEN(SUBSTITUTE(C2,"#", ""))
I	#+@/len	Calculates the % of # and @ signs in the text	=(G2+H2)/E2
J**	no. of " "	Counts the no. of spaces	=LEN(C2)-LEN(SUBSTITUTE(C2," ", ""))
K**	" "/len	Calculates the % of spaces	=J2/E2
L	no. of ?	Counts the no. of ?s	=LEN(C2)- LEN(SUBSTITUTE(C2,"?", ""))
M	?/len	Calculates the % of ?s	=L2/E2
N <sup>+</sup>	words/@+#	Identifies the tweets with most non-word chars	=(E2/J2)/(G2+H2)
O***	words1	Flags as TRUE tweets with listed words	See Appendix D, Words1
P***	ideograms and arabic	Flags as TRUE tweets with listed characters	See Appendix D, Ideograms and Arabic
Q***	diacritics	Flags as TRUE tweets with listed diacritics	See Appendix D, Diacritics
R	rand()	Generates a random no.	=RAND()
S	startoftweet	Date stamped start of text marker	startoftweet2021
T	concat	Joins the startoftweet to the tweet text	=CONCAT(S2," ",C2)
U		Counts displayed rows	=COUNT(D:D)

\* tweets with first 50 characters the same were deleted as duplicates

\*\* tweets using large numbers of spaces to create an image, were deleted

\*\*\* tweets flagged as TRUE were deleted

+ tweets with high numbers of accounts and hashtags were deleted

## Appendix D – Non-english language filtering formulae

### Words

```
=OR(  
ISNUMBER(SEARCH(" aku ",C5)), ISNUMBER(SEARCH(" anak ",C5)), ISNUMBER(SEARCH("  
auf ",C5)),  
ISNUMBER(SEARCH(" banyak ",C5)), ISNUMBER(SEARCH(" buku ",C5)),  
ISNUMBER(SEARCH(" cewe ",C5)),  
ISNUMBER(SEARCH(" da ",C5)), ISNUMBER(SEARCH(" de ",C5)), ISNUMBER(SEARCH(" des  
",C5)), ISNUMBER(SEARCH(" del ",C5)), ISNUMBER(SEARCH(" der ",C5)),  
ISNUMBER(SEARCH(" das ",C5)), ISNUMBER(SEARCH(" dass ",C5)), ISNUMBER(SEARCH(" der  
",C5)), ISNUMBER(SEARCH(" des ",C5)),  
ISNUMBER(SEARCH(" ein ",C5)), ISNUMBER(SEARCH(" eine ",C5)), ISNUMBER(SEARCH(" es  
",C5)), ISNUMBER(SEARCH(" est ",C5)), ISNUMBER(SEARCH(" este ",C5)),  
ISNUMBER(SEARCH(" ga ",C5)),  
ISNUMBER(SEARCH(" hai ",C5)), ISNUMBER(SEARCH(" haben ",C5)),  
ISNUMBER(SEARCH(" ini ",C5)), ISNUMBER(SEARCH(" iya ",C5)), ISNUMBER(SEARCH(" ich  
",C5)),  
ISNUMBER(SEARCH(" jg ",C5)),  
ISNUMBER(SEARCH(" kan ",C5)), ISNUMBER(SEARCH(" kalau ",C5)), ISNUMBER(SEARCH("  
kena ",C5)), ISNUMBER(SEARCH(" kok ",C5)), ISNUMBER(SEARCH(" ko ",C5)),  
ISNUMBER(SEARCH(" ku ",C5)),  
ISNUMBER(SEARCH(" la ",C5)), ISNUMBER(SEARCH(" le ",C5)), ISNUMBER(SEARCH(" les  
",C5)), ISNUMBER(SEARCH(" lakh ",C5)),  
ISNUMBER(SEARCH(" mana ",C5)), ISNUMBER(SEARCH(" mbak ",C5)),  
ISNUMBER(SEARCH(" na ",C5)), ISNUMBER(SEARCH(" ne ",C5)), ISNUMBER(SEARCH(" neit  
",C5)), ISNUMBER(SEARCH(" nie ",C5)), ISNUMBER(SEARCH(" nicht ",C5)),  
ISNUMBER(SEARCH(" nak ",C5)),  
ISNUMBER(SEARCH(" pun ",C5)),  
ISNUMBER(SEARCH(" que ",C5)), ISNUMBER(SEARCH(" que ",C5)),  
ISNUMBER(SEARCH(" se ",C5)), ISNUMBER(SEARCH(" si ",C5)), ISNUMBER(SEARCH(" som  
",C5)), ISNUMBER(SEARCH(" sih ",C5)), ISNUMBER(SEARCH(" suka ",C5)),  
ISNUMBER(SEARCH(" sein ",C5)),  
ISNUMBER(SEARCH(" tapi ",C5)),  
ISNUMBER(SEARCH(" un ",C5)), ISNUMBER(SEARCH(" une ",C5)), ISNUMBER(SEARCH(" und  
",C5)),  
ISNUMBER(SEARCH(" vor ",C5)), ISNUMBER(SEARCH(" voor ",C5)), ISNUMBER(SEARCH(" von  
",C5)),  
ISNUMBER(SEARCH(" werden ",C5)),  
ISNUMBER(SEARCH(" yg ",C5)),  
ISNUMBER(SEARCH(" ze ",C5)), ISNUMBER(SEARCH(" zu ",C5)), ISNUMBER(SEARCH(" ze  
",C5)),  
)
```

## Ideograms and Arabic

=OR( ISNUMBER(SEARCH("人",C5)), ISNUMBER(SEARCH("中",C5)), ISNUMBER(SEARCH("会",C5)), ISNUMBER(SEARCH("生",C5)), ISNUMBER(SEARCH("子",C5)), ISNUMBER(SEARCH("今",C5)), ISNUMBER(SEARCH("時",C5)), ISNUMBER(SEARCH("見",C5)), ISNUMBER(SEARCH("市",C5)), ISNUMBER(SEARCH("力",C5)), ISNUMBER(SEARCH("自",C5)), ISNUMBER(SEARCH("前",C5)), ISNUMBER(SEARCH("者",C5)), ISNUMBER(SEARCH("体",C5)), ISNUMBER(SEARCH("女",C5)), ISNUMBER(SEARCH("言",C5)), ISNUMBER(SEARCH("思",C5)), ISNUMBER(SEARCH("实",C5)), ISNUMBER(SEARCH("問",C5)), ISNUMBER(SEARCH("話",C5)), ISNUMBER(SEARCH("政",C5)), ISNUMBER(SEARCH("意",C5)), ISNUMBER(SEARCH("党",C5)), ISNUMBER(SEARCH("的",C5)), ISNUMBER(SEARCH("一",C5)), ISNUMBER(SEARCH("了",C5)), ISNUMBER(SEARCH("是",C5)), ISNUMBER(SEARCH("我",C5)), ISNUMBER(SEARCH("不",C5)), ISNUMBER(SEARCH("在",C5)), ISNUMBER(SEARCH("人",C5)), ISNUMBER(SEARCH("们",C5)), ISNUMBER(SEARCH("有",C5)), ISNUMBER(SEARCH("것",C5)), ISNUMBER(SEARCH("하",C5)), ISNUMBER(SEARCH("있",C5)), ISNUMBER(SEARCH("수",C5)), ISNUMBER(SEARCH("나",C5)), ISNUMBER(SEARCH("없",C5)), ISNUMBER(SEARCH("않",C5)), ISNUMBER(SEARCH("사",C5)), ISNUMBER(SEARCH("우",C5)), ISNUMBER(SEARCH("Г",C5)), ISNUMBER(SEARCH("Д",C5)), ISNUMBER(SEARCH("Ж",C5)), ISNUMBER(SEARCH("И",C5)), ISNUMBER(SEARCH("Й",C5)), ISNUMBER(SEARCH("К",C5)), ISNUMBER(SEARCH("Л",C5)), ISNUMBER(SEARCH("П",C5)), ISNUMBER(SEARCH("Ф",C5)), ISNUMBER(SEARCH("Ц",C5)), ISNUMBER(SEARCH("Ч",C5)), ISNUMBER(SEARCH("Ш",C5)), ISNUMBER(SEARCH("Щ",C5)), ISNUMBER(SEARCH("Ъ",C5)), ISNUMBER(SEARCH("Ы",C5)), ISNUMBER(SEARCH("Э",C5)), ISNUMBER(SEARCH("Ю",C5)), ISNUMBER(SEARCH("Я",C5)), ISNUMBER(SEARCH("б",C5)), ISNUMBER(SEARCH("г",C5)), ISNUMBER(SEARCH("д",C5)), ISNUMBER(SEARCH("ж",C5)), ISNUMBER(SEARCH("и",C5)), ISNUMBER(SEARCH("й",C5)), ISNUMBER(SEARCH("к",C5)), ISNUMBER(SEARCH("л",C5)), ISNUMBER(SEARCH("п",C5)), ISNUMBER(SEARCH("ф",C5)), ISNUMBER(SEARCH("ц",C5)), ISNUMBER(SEARCH("ч",C5)), ISNUMBER(SEARCH("ш",C5)), ISNUMBER(SEARCH("щ",C5)), ISNUMBER(SEARCH("э",C5)), ISNUMBER(SEARCH("ю",C5)), ISNUMBER(SEARCH("я",C5)), ISNUMBER(SEARCH("ء",C5)), ISNUMBER(SEARCH("ؤ",C5)), ISNUMBER(SEARCH("ئ",C5)), ISNUMBER(SEARCH("آ",C5)), ISNUMBER(SEARCH("إ",C5)), ISNUMBER(SEARCH("أ",C5)), ISNUMBER(SEARCH("ب",C5)), ISNUMBER(SEARCH("ة",C5)), ISNUMBER(SEARCH("ت",C5)), ISNUMBER(SEARCH("ث",C5)), ISNUMBER(SEARCH("ج",C5)), ISNUMBER(SEARCH("ح",C5)), ISNUMBER(SEARCH("خ",C5)), ISNUMBER(SEARCH("د",C5)), ISNUMBER(SEARCH("ذ",C5)), ISNUMBER(SEARCH("ر",C5)), ISNUMBER(SEARCH("ز",C5)), ISNUMBER(SEARCH("س",C5)), ISNUMBER(SEARCH("ش",C5)), ISNUMBER(SEARCH("ص",C5)), ISNUMBER(SEARCH("ض",C5)), ISNUMBER(SEARCH("ط",C5)), ISNUMBER(SEARCH("ظ",C5)), ISNUMBER(SEARCH("ع",C5)), ISNUMBER(SEARCH("غ",C5)), ISNUMBER(SEARCH("ف",C5)), ISNUMBER(SEARCH("ق",C5)), ISNUMBER(SEARCH("ك",C5)), ISNUMBER(SEARCH("ل",C5)), ISNUMBER(SEARCH("م",C5)), ISNUMBER(SEARCH("ن",C5)), ISNUMBER(SEARCH("ه",C5)), ISNUMBER(SEARCH("و",C5)), ISNUMBER(SEARCH("ى",C5)), ISNUMBER(SEARCH("ي",C5)), ISNUMBER(SEARCH("ي",C5)), ISNUMBER(SEARCH("पै",C5)), ISNUMBER(SEARCH("है",C5)), ISNUMBER(SEARCH("हैं",C5)),

## Sexual violence and the gender pay gap

ISNUMBER(SEARCH("नहीं",C5)), ISNUMBER(SEARCH("लिए",C5)),  
ISNUMBER(SEARCH("गया",C5)), ISNUMBER(SEARCH("तथा",C5)),  
ISNUMBER(SEARCH("अपने",C5)), ISNUMBER(SEARCH("कुछ",C5)),  
ISNUMBER(SEARCH("साथ",C5)), ISNUMBER(SEARCH("होता",C5)),  
ISNUMBER(SEARCH("था",C5)), ISNUMBER(SEARCH("दिया",C5)),  
ISNUMBER(SEARCH("हुए",C5)), ISNUMBER(SEARCH("कोई",C5)),  
ISNUMBER(SEARCH("रूप",C5)), ISNUMBER(SEARCH("से",C5)), ISNUMBER(SEARCH("मैं",C5)),  
ISNUMBER(SEARCH("रहा",C5))  
)

**Diacritics**

```

=OR(
ISNUMBER(SEARCH("ä",C5)), ISNUMBER(SEARCH("å",C5)), ISNUMBER(SEARCH("á",C5)),
ISNUMBER(SEARCH("ã",C5)), ISNUMBER(SEARCH("ā",C5)), ISNUMBER(SEARCH("À",C5)),
ISNUMBER(SEARCH("à",C5)), ISNUMBER(SEARCH("Ä",C5)), ISNUMBER(SEARCH("ǎ",C5)),
ISNUMBER(SEARCH("ɑ",C5)),
ISNUMBER(SEARCH("ß",C5)), ISNUMBER(SEARCH("β",C5)),
ISNUMBER(SEARCH("č",C5)), ISNUMBER(SEARCH("ç",C5)), ISNUMBER(SEARCH("ć",C5)),
ISNUMBER(SEARCH("ḍ",C5)),
ISNUMBER(SEARCH("è",C5)), ISNUMBER(SEARCH("é",C5)), ISNUMBER(SEARCH("è",C5)),
ISNUMBER(SEARCH("ë",C5)), ISNUMBER(SEARCH("ē",C5)), ISNUMBER(SEARCH("Ě",C5)),
ISNUMBER(SEARCH("è",C5)), ISNUMBER(SEARCH("ě",C5)),
ISNUMBER(SEARCH("ƒ",C5)),
ISNUMBER(SEARCH("ğ",C5)),
ISNUMBER(SEARCH("ħ",C5)), ISNUMBER(SEARCH("μ",C5)),
ISNUMBER(SEARCH("ı",C5)),
ISNUMBER(SEARCH("ĵ",C5)),
ISNUMBER(SEARCH("ķ",C5)),
ISNUMBER(SEARCH("ï",C5)),
ISNUMBER(SEARCH("ḿ",C5)),
ISNUMBER(SEARCH("ñ",C5)), ISNUMBER(SEARCH("ń",C5)), ISNUMBER(SEARCH("ņ",C5)),
ISNUMBER(SEARCH("и",C5)), ISNUMBER(SEARCH("л",C5)),
ISNUMBER(SEARCH("ö",C5)), ISNUMBER(SEARCH("ô",C5)), ISNUMBER(SEARCH("ó",C5)),
ISNUMBER(SEARCH("ø",C5)), ISNUMBER(SEARCH("œ",C5)),
ISNUMBER(SEARCH("Ṗ",C5)),
ISNUMBER(SEARCH("ḻ",C5)),
ISNUMBER(SEARCH("ř",C5)), ISNUMBER(SEARCH("ŕ",C5)), ISNUMBER(SEARCH("я",C5)),
ISNUMBER(SEARCH("ș",C5)), ISNUMBER(SEARCH("š",C5)), ISNUMBER(SEARCH("ś",C5)),
ISNUMBER(SEARCH("ş",C5)),
ISNUMBER(SEARCH("т",C5)), ISNUMBER(SEARCH("т",C5)),
ISNUMBER(SEARCH("ü",C5)), ISNUMBER(SEARCH("û",C5)), ISNUMBER(SEARCH("ū",C5)),
ISNUMBER(SEARCH("ú",C5)), ISNUMBER(SEARCH("Û",C5)),
ISNUMBER(SEARCH("ṽ",C5)),
ISNUMBER(SEARCH("Ẁ",C5)),
ISNUMBER(SEARCH("ж",C5)),
ISNUMBER(SEARCH("ÿ",C5)), ISNUMBER(SEARCH("λ",C5)), ISNUMBER(SEARCH("ÿ",C5)),
ISNUMBER(SEARCH("ž",C5)), ISNUMBER(SEARCH("ž",C5)),
ISNUMBER(SEARCH("ı",C5)), ISNUMBER(SEARCH("ı",C5)), ISNUMBER(SEARCH("ю",C5)),
ISNUMBER(SEARCH("ч",C5)), ISNUMBER(SEARCH(" П",C5)), ISNUMBER(SEARCH("д",C5)),
ISNUMBER(SEARCH("з",C5)),
ISNUMBER(SEARCH("ǰ",C5)), ISNUMBER(SEARCH("ǰ",C5)), ISNUMBER(SEARCH("J",C5)),
ISNUMBER(SEARCH("ن",C5)), ISNUMBER(SEARCH("ط",C5)), ISNUMBER(SEARCH("ن",C5)),
ISNUMBER(SEARCH("ن",C5)),
)

```

**Appendix E - Nouns modifying FEMALE**

The industries referred to below are taken from the Statistica chart (Figure 30), supplemented by the category \*Executive & Managerial which cuts across all industries

**TwitterPG corpus**

Noun	Frequency	LogDice	Industry
doctor	52	9.6	Health & social work
athlete	42	9.4	Arts, entertainment & recreation
professor	36	9.2	Education
player	44	9.1	Arts, entertainment & recreation
leader	31	8.9	*Executive & Managerial
CEO	31	8.8	*Executive & Managerial
physician	25	8.7	Health & social work
model	22	8.4	Arts, entertainment & recreation
graduate	19	8.2	Other services
firefighter	17	8.2	Other services
footballer	15	8	Arts, entertainment & recreation
engineer	15	8	Professional, scientific & technical
sport	15	7.8	Arts, entertainment & recreation
nurse	14	7.8	Health & social work
surgeon	13	7.8	Health & social work
executive	13	7.7	*Executive & Managerial
artist	12	7.7	Arts, entertainment & recreation
actor	12	7.6	Arts, entertainment & recreation
MP	11	7.5	Public administration & defence
writer [literature]	11	7.5	Arts, entertainment & recreation
teacher	12	7.4	Education
freelancer	10	7.4	Arts, entertainment & recreation
entrepreneur	10	7.4	*Executive & Managerial
founder	10	7.4	*Executive & Managerial
presenter	9	7.2	Information & Communication
barrister	8	7.1	Professional, scientific & technical
specialist	8	7.1	Health & social work
staffer [government]	8	7.1	Public administration & defence
influencer	8	7	Information & Communication
marketer	7	6.9	Information & Communication

Noun	Frequency	LogDice	Industry
streamer	7	6.9	Information & Communication
journalist	7	6.9	Information & Communication
academic	7	6.9	Education
president	7	6.9	Public administration & defence
chair	6	6.7	*Executive & Managerial
economist	6	6.7	Professional, scientific & technical
lawyer	6	6.7	Professional, scientific & technical
scientist	6	6.7	Professional, scientific & technical
prof	5	6.4	Education
football	5	6.4	Arts, entertainment & recreation
inventor	4	6.1	Professional, scientific & technical
pilot	4	6.1	Transport & storage
celebrity	4	6.1	Arts, entertainment & recreation
GP	4	6.1	Health & social work
minister	4	6.1	Public administration & defence
star	4	6	Arts, entertainment & recreation
boss	4	6	*Executive & Managerial
driver	4	6	Transport & storage
Physicians	3	5.7	Health & social work
pornstar	3	5.7	Arts, entertainment & recreation
cricketer	3	5.7	Arts, entertainment & recreation
superstar	3	5.7	Arts, entertainment & recreation
creative	3	5.7	Arts, entertainment & recreation
secretary	3	5.7	Administrative & support services
adviser	3	5.7	Professional, scientific & technical
cleaner	3	5.7	Accommodation & food services
fighter [fire]	3	5.7	Other services
faculty	3	5.7	Education
actress	3	5.7	Arts, entertainment & recreation
coach	3	5.7	Education

## Sexual violence and the gender pay gap

### NexisPG corpus

Noun	Frequency	LogDice	Industry
graduate	90	10	Other services
doctor	55	9.4	Health & social work
candidate	42	8.9	Other services
CEO	39	8.9	*Executive & Managerial
leader	40	8.7	*Executive & Managerial
player	32	8.6	Arts, entertainment & recreation
executive	34	8.4	*Executive & Managerial
presenter	27	8.4	Information & Communication
athlete	25	8.4	Arts, entertainment & recreation
star	27	8.3	Arts, entertainment & recreation
manager	28	8.3	*Executive & Managerial
barrister	22	8.2	Professional, scientific & technical
entrepreneur	22	8.2	*Executive & Managerial
partner	23	8.1	*Executive & Managerial
director	27	8	*Executive & Managerial
physician	19	7.9	Health & social work
professor	19	7.9	Education
lawyer	17	7.8	Professional, scientific & technical
staffer [government]	16	7.7	Public administration & defence
boss	16	7.6	*Executive & Managerial
actor	15	7.6	Arts, entertainment & recreation
minister	17	7.5	Public administration & defence
owner	12	7.2	*Executive & Managerial
artist	11	7.2	Arts, entertainment & recreation
engineer	11	7.2	Professional, scientific & technical
MP	12	7.2	Public administration & defence
surgeon	10	7	Health & social work
president	9	6.7	Public administration & defence
scientist	8	6.7	Professional, scientific & technical
GP	7	6.5	Health & social work
officer	8	6.5	Arts, entertainment & recreation
nurse	7	6.4	Health & social work
pilot	6	6.3	Transport & storage

<b>Noun</b>	<b>Frequency</b>	<b>LogDice</b>	<b>Industry</b>
academic	6	6.3	Education
journalist	6	6.3	Information & Communication
model	6	6.2	Arts, entertainment & recreation
pediatrician	5	6.1	Health & social work
golfer	5	6.1	Arts, entertainment & recreation
mayor	5	6.1	Public administration & defence
influencer	5	6	Information & Communication
founder	5	6	*Executive & Managerial
teacher	5	5.9	Education
priest	4	5.8	Other services
cricketer	4	5.7	Arts, entertainment & recreation
apprentice	4	5.7	Other services
CFO	4	5.7	*Executive & Managerial
co-star	4	5.7	Arts, entertainment & recreation
coach	4	5.7	Education
lawmaker	4	5.7	Public administration & defence
lead	4	5.7	*Executive & Managerial
chair	4	5.7	*Executive & Managerial
head	4	5.6	*Executive & Managerial

## References

- @katieaka. (2016, May 10). *New research: The value of influencers on Twitter*. Twitter.  
[https://blog.twitter.com/en\\_us/a/2016/new-research-the-value-of-influencers-on-twitter.html](https://blog.twitter.com/en_us/a/2016/new-research-the-value-of-influencers-on-twitter.html)
- Advance Pro Bono. (2021). *Prevalence and reporting of sexual harassment in UK public spaces*. 'APPG for UN Women'.  
<https://publications.parliament.uk/pa/cm/cmllparty/rules/guide.html>
- Ahrens, C. (2006). Being silenced: the impact of negative social reactions on the disclosure of rape. *American journal of community psychology*, 38(3-4), 263-274.  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1705531/>
- Alkhamash, R. (2021). The social media framing of gender pay gap debate in American women's sport: A linguistic analysis of emotive language. *Training, language and culture (Online)*, 5(4), 22-35. <https://doi.org/10.22363/2521-442X-2021-5-4-22-35>
- Alonso, O., Fetterly, D., & Manasse, M. (2013). *Duplicate News Story Detection Revisited* (Technical Report MSR-TR-2013-60). Microsoft. <https://www.microsoft.com/en-us/research/wp-content/uploads/2013/05/paper-1.pdf>
- Amadjarif, Z., Angeli, M., Haldane, A., & Zemaityte, G. (2020). Understanding pay gaps [Staff Working Paper No. 877](877). Bank of England.  
<https://www.bankofengland.co.uk/working-paper/2020/understanding-pay-gaps>
- American Association of University Women. (2015). *Deeper in debt, women and student loans*. A. A. o. U. Women.  
<https://www.aauw.org/app/uploads/2020/03/DeeperinDebt-nsa.pdf>
- Ammerman, C., & Groysberg, B. (2017). Why Sexual Harassment Persists and What Organizations Can Do to Stop It. *Harvard Business Review*.  
<https://hbr.org/2017/12/why-sexual-harassment-persists-and-what-organizations-can-do-to-stop-it>
- Amnesty International. (2018). *Toxic Twitter – A Toxic Place for Women*.  
<https://www.amnesty.org/en/latest/research/2018/03/online-violence-against-women-chapter-1-1/>
- Anderson, D. (2017). Conceptual competence injustice. *Social Epistemology*, 31(2), 210-223.  
<https://doi.org/10.1080/02691728.2016.1241320>
- Anderson, D. (2020). Linguistic Hijacking. *Feminist philosophy quarterly*, 6(3).  
<https://doi.org/10.5206/fpq/2020.3.8162>
- Anderson, D. (2021). *Metasemantics and Intersectionality in the Misinformation Age : Truth in Political Struggle* (1st ed.). Springer International Publishing AG.  
<https://ebookcentral.proquest.com/lib/lancaster/reader.action?docID=6719980&ppg=19#>
- Andrew, A., Bandiera, O., Costa Dias, M., & Landais, C. (2024). Women and men at work. *Oxford Open Economics*, 3 (Supplement\_1, Dimensions of Inequality: The IFS Deaton

- Review), i294-i322.  
[https://academic.oup.com/ooec/article/3/Supplement\\_1/i294/7708109](https://academic.oup.com/ooec/article/3/Supplement_1/i294/7708109)
- Angouri, J., & Tseliga, T. (2010). "You Have No Idea What You are Talking About!" From e-disagreement to e-impoliteness in two online fora. *Journal of politeness research : language, behaviour, culture*, 6(1), 57-82. <https://www-degruyterbrill-com.ezproxy.lancs.ac.uk/document/doi/10.1515/jplr.2010.004/html>
- Auspurg, K., Hinz, T., & Sauer, C. (2017). Why Should Women Get Less? Evidence on the Gender Pay Gap from Multifactorial Survey Experiments. *Am Sociol Rev*, 82(1), 179-210. <https://doi.org/10.1177/0003122416683393>
- Austin, J. L. (1975). *How To Do Things With Words: The William James Lectures delivered at Harvard University in 1955*. Oxford University Press.  
<https://doi.org/10.1093/acprof:oso/9780198245537.001.0001>
- Badenhausen, K. (2024). Women Shut Out From Highest-Paid Athletes Despite Investment Boom. *Sportico*.  
<https://www.proquest.com/docview/2923124449?accountid=11979&parentSessionId=Md6sJ1pgJYUOUiJPypHCUGDkcHAMni6lkg4Dcq7COCg%3D&pq-origsite=primo&sourcetype=Other%20Sources>
- Bailey, R., Dugard, J., Smith, S., & Porges, S. (2023). Appeasement: replacing Stockholm syndrome as a definition of a survival strategy. *European journal of psychotraumatology*, 14(1), 2161038.  
<https://www.tandfonline.com/doi/pdf/10.1080/20008066.2022.2161038>
- Baker, P. (2009). The BE06 Corpus of British English and recent language change. *International Journal of Corpus Linguistics*, 14(3), 312-337. <http://www.jbe-platform.com/content/journals/10.1075/ijcl.14.3.02bak>
- Baker, P. (2010). *Sociolinguistics and corpus linguistics*. Edinburgh University Press.  
<https://ebookcentral.proquest.com/lib/lancaster/detail.action?pq-origsite=primo&docID=536982>
- Baker, P. (2014). *Using corpora to analyze gender*. Bloomsbury Academic.
- Baker, P., Gabrielatos, C., Khosravini, M., Krzyżanowski, M., McEnery, A., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society*, 19(3), 273-306.  
<https://doi.org/10.1177/0957926508088962>
- Baker, P., & Levon, E. (2015). Picking the right cherries? A comparison of corpus-based and qualitative analyses of news articles about masculinity. *Discourse & communication*, 9(2), 221-236. <https://doi.org/10.1177/1750481314568542>
- Baker, P., & McEnery, A. (2015). *Corpora and Discourse Studies, Integrating Discourse and Corpora*. Palgrave Macmillan Limited.  
<https://ebookcentral.proquest.com/lib/lancaster/detail.action?docID=4001303>
- Banet-Weiser, S. (2019). Radical vulnerability: Feminism, victimhood and agency. In M. Gámez Fuentes, S. Núñez Puente, & E. Gómez Nicolau (Eds.), *Re-writing Women as*

- Victims* (1st ed., pp. 167-180). Routledge. <https://doi.org/10.4324/9781351043601-13>
- Barca, L. (2018). The agency factor: neoliberal configurations of risk in news discourse on the Steubenville, Ohio rape case. *Critical discourse studies*, 15(3), 265-284. <https://doi.org/10.1080/17405904.2017.1408476>
- Barry, E. (2018). Happy 'National Jealousy Day'! Finland Bares Its Citizens' Taxes. *New York Times*. <https://www.nytimes.com/2018/11/01/world/europe/finland-national-jealousy-day.html>
- Baxter, J. (2003). *Positioning gender in discourse: a feminist methodology*. Palgrave Macmillan.
- BBC. (2014). *Two guilty over abusive tweets to Caroline Criado-Perez*. <https://www.bbc.co.uk/news/uk-25641941>
- BBC. (2018a). *BBC China editor Carrie Gracie quits post in equal pay row*. <https://www.bbc.com/news/uk-42598775>
- BBC. (2018b). *Christine Blasey Ford: Kavanaugh sex attack changed my life*. <https://www.bbc.com/news/world-us-canada-45664863>
- BBC. (2019). *Bollywood actor Nana Patekar cleared of sexual harassment charges*. <https://www.bbc.co.uk/news/world-asia-india-48593301>
- Ben-David, S. (2020). From Victim to Survivor to Overcomer. In J. Joseph & S. Jergenson (Eds.), *An International Perspective on Contemporary Developments in Victimology*. (pp. 21-30). Springer. [https://doi.org/https://doi.org/10.1007/978-3-030-41622-5\\_2](https://doi.org/https://doi.org/10.1007/978-3-030-41622-5_2)
- Bennett, F. (2024). Gendered economic inequalities: a social policy perspective. *Oxford Open Economics*, 3 (Supplement\_1, Dimensions of Inequality: The IFS Deaton Review), i345-i353. <https://doi.org/10.1093/ooec/odad030>
- Bilenka, Y., & Kolomiets, M. (2021). Features of Youth Communication and the Usage of Discourse Markers 32(71). *Vernadsky Journals* [https://www.philol.vernadskyjournals.in.ua/journals/2021/4\\_2021/part\\_1/17.pdf](https://www.philol.vernadskyjournals.in.ua/journals/2021/4_2021/part_1/17.pdf)
- Bouvier, G. (2022). From 'echo chambers' to 'chaos chambers': discursive coherence and contradiction in the #MeToo Twitter feed. *Critical discourse studies*, 19(2), 179-195. <https://doi.org/10.1080/17405904.2020.1822898>
- Brezina, V. (2018). Statistical Choices in Corpus-based Discourse Analysis. In C. Taylor & A. Marchi (Eds.). *Corpus approaches to discourse : a critical review*. Routledge. <https://ebookcentral.proquest.com/lib/lancaster/detail.action?docID=5254458>
- Britannica [electronic resource], (2023). Gone too far. In *britannica.com*. <https://www.britannica.com/topic/political-correctness>
- Buchanan, I. (2025). *Women in politics and public life* (01250). H. o. C. Library. chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://researchbriefings.files.parliament.uk/documents/SN01250/SN01250.pdf
- Business of Apps. (2020, August 18). *Twitter Revenue and Usage Statistics (2020)*. <https://www.businessofapps.com/data/twitter-statistics/>

- Butler, D. (2020). *Speech on International Women's Day* [Speech]. [www.ukpol.co.uk/dawn-butler-2020-speech-on-international-womens-day/](http://www.ukpol.co.uk/dawn-butler-2020-speech-on-international-womens-day/)
- Butler, J. (1999). *Gender Trouble: Feminism and the Subversion of Identity*. Taylor & Francis. <https://doi-org.ezproxy.lancs.ac.uk/10.4324/9780203902752>
- Butler, S. (2025). Thousands of Asda workers win latest stage in decade-long equal pay case. *The Guardian*. <https://www.theguardian.com/business/2025/feb/03/thousands-asda-workers-win-latest-stage-equal-pay-case>
- Cambridge Dictionary [electronic resource]. (n.d.-a). Fall over yourself. In [dictionary.cambridge.org](https://dictionary.cambridge.org). <https://dictionary.cambridge.org/dictionary/english/fall-over-self>
- Cambridge Dictionary [electronic resource]. (n.d.-b). Right wing. In [dictionary.cambridge.org](https://dictionary.cambridge.org). <https://dictionary.cambridge.org/dictionary/english/right-wing>
- Campoamor, D. (2018, May 21). I'm Not a Sexual Assault "Survivor"—I'm a Victim. *Harpers Bazaar*. <https://www.harpersbazaar.com/culture/features/a20138398/stop-using-survivor-to-describe-sexual-assault-victims/>
- Carpenter, J. (2018). The link between the gender wage gap and sexual harassment. *CNN Money*. <https://money.cnn.com/2018/01/04/pf/pay-gap-sexual-harassment/index.html>
- Cheng, S., Corrington, A., Hebl, M., Ng, L., & Watson, I. (2018). Victim Precipitation and the Wage Gap. *Ind. Organ. Psychol*, 11(1), 144-151. <https://doi.org/10.1017/iop.2017.100>
- Chouliaraki, L. (2024). *Wronged : The Weaponization of Victimhood* (1st ed.). Columbia University Press. <https://ebookcentral.proquest.com/lib/lancaster/reader.action?docID=7373840>
- Christie, N. (1986). The ideal victim. In E. A. Fattah (Ed.), *From Crime Policy to Victim Policy* (pp. 17-30). Palgrave Macmillan. <https://doi.org/10.1007/978-1-349-08305-3>
- Citizen's Advice Bureau. (n.d.). *Check what it might cost to make an employment tribunal claim*. <https://www.citizensadvice.org.uk/work/employment-tribunal/employment-tribunals/before-you-go-to-the-tribunal/check-what-it-might-cost-to-make-an-employment-tribunal-claim/>
- Clark, D. (2023a). *Gender pay gap for gross hourly earnings in the United Kingdom in 2023, by industry sector*. Statista. <https://www.statista.com/statistics/760342/gender-pay-gap-uk-by-sector/>
- Clark, D. (2023b). *Number of employees on a zero-hours contract in the United Kingdom from 2013 to 2023, by gender*. Statista. <https://www.statista.com/statistics/398576/number-of-employees-zero-hour-contracts-gender/>
- Clayton, I. (2025). *Fawning : Why the Need to Please Makes Us Lose Ourselves - and How to Find Our Way Back* (1st ed.). LEAP. <https://ebookcentral.proquest.com/lib/lancaster/reader.action?docID=31680724&c=RVBVQg&ppg=1>

- Clement, J. (2019a, July 24). *Countries with the most Twitter users 2020*. Statista. <https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/>
- Clement, J. (2019b, July 16). *Most popular mobile social networking apps in the United States as of September 2019, by average session duration*. Statista. <https://www.statista.com/statistics/579411/top-us-social-networking-apps-ranked-by-session-length/#statisticContainer>
- Cole, A. (2007). *The cult of true victimhood : from the war on welfare to the war on terror*. Stanford University Press.
- Cole, A. (2021). (Re)Made in America: Survivorship after the Shoah. *European journal of cultural studies*, 24(1), 28-44. <https://journals.sagepub.com/doi/epub/10.1177/1367549420985855>
- Collins Dictionary. (n.d.-a). Oh my gawd. In *collinsdictionary.com [electronic resource]* <https://www.collinsdictionary.com/dictionary/english/gawd>
- Collins Dictionary. (n.d.-b). Victim. In *collinsdictionary.com [electronic resource]*. <https://www.collinsdictionary.com/dictionary/english/victim>
- Collinson, P. (2016). Norway, the country where you can see everyone's tax returns. *The Guardian*. <https://www.theguardian.com/money/blog/2016/apr/11/when-it-comes-to-tax-transparency-norway-leads-the-field>
- Conaghan, J. (2013). *Law and Gender*. Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780199592920.001.0001>
- Consortium, B.N.C. (2022). *What is the British National Corpus?* <http://www.natcorp.ox.ac.uk/>
- Cooke, L. (2024). Gender revolution, evolution or neverlution? *Oxford Open Economics*, 3(Supplement\_1, Dimensions of Inequality: The IFS Deaton Review), i354-i361. <https://doi.org/10.1093/ooec/odad011>
- Coulthard, M., & Johnson, A. (2007). *An Introduction to forensic linguistics : language in evidence*. Routledge. <https://ebookcentral.proquest.com/lib/lancaster/detail.action?docID=325298>
- Council of Europe. (n.d.). *The European Convention on Human Rights and its protocols*. [https://www.echr.coe.int/documents/d/echr/convention\\_eng](https://www.echr.coe.int/documents/d/echr/convention_eng)
- Cowan, S. (2019). Sense and Sensibilities: A Feminist Critique of Legal Interventions against Sexual Violence. *The Edinburgh law review*, 23(1), 22-51. <https://www-eupublishing-com.ezproxy.lancs.ac.uk/doi/epub/10.3366/elr.2019.0523>
- Cox, C. (2020). 'Healthcare Heroes': problems with media focus on heroism from healthcare workers during the COVID-19 pandemic. *Journal of medical ethics*, 46(8), 510-513. <https://jme-bmj-com.ezproxy.lancs.ac.uk/content/46/8/510>
- Criado-Perez, C. (2019). *Invisible women : exposing data bias in a world designed for men*. Chatto & Windus

- Crown Prosecution Service. (2021, November). *Rape and Sexual Offences - Annex A: Tackling Rape Myths and Stereotypes*. <https://www.cps.gov.uk/legal-guidance/rape-and-sexual-offences-annex-tackling-rape-myths-and-stereotypes>
- Crown Prosecution Service. (n.d.). *Rape and Sexual Offences - Chapter 21*. <https://www.cps.gov.uk/legal-guidance/rape-and-sexual-offences-chapter-21-societal-myths>
- Crystal, D. (2006). *Language and the Internet* (Second ed.). Cambridge University Press.
- De Sacco, E. (2019). Survivor speech: Formulations of the victim/survivor identity in contemporary UK media. In M. J. Gámez Fuentes, S. Núñez Puente, & E. Gómez Nicolau (Eds.), *Re-writing Women as Victims: From Theory to Practice* (1st ed., pp. 194-209). Routledge. <https://doi.org/10.4324/9781351043601-15>
- De Simone, D. (2024). Met only sought Al Fayed charges for two victims. *BBC*. <https://www.bbc.com/news/articles/cp8egz8gr8lo>
- Delker, B., Salton, R., McLean, K., & Syed, M. (2020). Who has to tell their trauma story and how hard will it be? Influence of cultural stigma and narrative redemption on the storying of sexual violence. *PLoS one*, 15(6). <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7274398/>
- Digital Marketing Institute. (2020, August 15). *20 Influencer Marketing Statistics that will Surprise You*. <https://digitalmarketinginstitute.com/blog/20-influencer-marketing-statistics-that-will-surprise-you>
- Dillenburg, K. (2008). A behavior analytic perspective on victimology. *The Journal of Behavior Analysis of Offender and Victim Treatment and Prevention*, 1(1), 5-19. <https://web-p-ebsohost-com.ezproxy.lancs.ac.uk/ehost/detail/detail?vid=5&sid=94c84ac5-94d5-4a39-b4f4-518d73b14537%40redis&bdata=JnNpdGU9ZWVhc3QtbGl2ZQ%3d%3d#AN=2014-55084-002&db=pdh>
- Dominus, S. (2020). Breaking the Salary Sharing Taboo. *New York Times*. <https://www.nytimes.com/interactive/2020/02/19/magazine/salary-sharing.html>
- Duguid, A., & Partington, A. (2018). Absence: You don't know what you're missing. Or do you? In C. Taylor & A. Marchi (Eds.), *Corpus approaches to discourse : a critical review*. Routledge. <https://ebookcentral.proquest.com/lib/lancaster/detail.action?docID=5254458>
- Duncan, A., & Holcombe, J. (2021). *Over a quarter of a century until gender pay gap likely to close*. Bankwest Curtin Economics Centre. <https://bcec.edu.au/media/over-a-quarter-of-a-century-until-gender-pay-gap-likely-to-close/>
- Dunn, J. (2008). Accounting for Victimization: Social Constructionist Perspectives. *Sociology compass*, 2(5), 1601-1620. <https://doi.org/10.1111/j.1751-9020.2008.00150.x>
- Ehrlich, S. (2001). *Representing Rape : Language and Sexual Consent*. Routledge. <https://web-p-ebsohost-com.ezproxy.lancs.ac.uk/ehost/ebookviewer/ebook/bmxlYmtfXzkyMzg0X19BTg2?sid=a99e604f-3e93-4a50-90a6-8a7421aab274@redis&vid=0&format=EB&rid=1>

- Equality Now. (2017). *The World's Shame: The Global Rape Epidemic*.  
[https://equalitynow.storage.googleapis.com/wp-content/uploads/2017/11/16112622/EqualityNowRapeLawReport2017\\_Single\\_Pages\\_0-compressed.pdf](https://equalitynow.storage.googleapis.com/wp-content/uploads/2017/11/16112622/EqualityNowRapeLawReport2017_Single_Pages_0-compressed.pdf)
- Fairclough, N. (2001). *Language and Power* (Second Edition ed.). Routledge.
- Fairclough, N. (2009). Language, Reality and Power. In J. Culpeper, F. Katamba, P. Kerswill, R. Wodak, & A. McEnery (Eds.), *English Language: Description, Variation and Context*. Palgrave.
- Favaro, L., & Gill, R. (2019). Pump up the positivity': Neoliberalism, affective entrepreneurship and the victimhood/agency debate. In M. Gámez Fuentes, S. Núñez Puente, & E. Gómez Nicolau (Eds.), *Re-writing Women as Victims: From Theory to Practice* (1st ed., pp. 153-166). Routledge. <https://doi.org/10.4324/9781351043601-12>
- Feeney, N. (2015). Jennifer Lawrence: Playing Hunger Games' Katniss Inspired Me to Fight Pay Gap. *Time*. <https://time.com/4099494/jennifer-lawrence-gender-pay-gap-katniss-hunger-games/>
- Fineman, M. (1995). *The neutered mother, the sexual family, and other twentieth century tragedies*. Routledge. <https://doi.org/10.4324/9781315021744>
- Finkel, E., Bail, C., Cikara, M., Ditto, P., Iyengar, S., Klar, S., Druckman, J. (2020). Political sectarianism in America. *Science (American Association for the Advancement of Science)*, 370(6516), 533-536.  
<https://www.science.org/doi/10.1126/science.abe1715>
- Finkelstein, K. (2021). Year After Brick Attack, Refusing to 'Tiptoe Around Life'. *New York Times*, 25. <https://www.nytimes.com/2000/12/24/nyregion/year-after-brick-attack-refusing-to-tiptoe-around-life.html>
- Fisher, J., & Jones, E. (2024). The problem with resilience. *International journal of mental health nursing*, 33(1), 185-188. [onlinelibrary.wiley.com/doi/10.1111/inm.13220](https://onlinelibrary.wiley.com/doi/10.1111/inm.13220)
- flyflyer. (2022). Why Are Women's Tennis Matches 3 Sets? *My Tennis Outfitter*.  
<https://mytennisoutfitter.com/why-are-womens-tennis-matches-3-sets/>
- Fohring, S. (2012). *The Process of Victimisation: Investigating Risk, Reporting and Service Use* [Doctoral Thesis, University of Edinburgh], University of Edinburgh.  
<https://era.ed.ac.uk/bitstream/handle/1842/6436/Fohring2012.pdf?sequence=1&isAllowed=y>
- Fohring, S. (2015). An integrated model of victimization as an explanation of non-involvement with the criminal justice system. *International review of victimology*, 21(1), 45-70. <https://doi.org/10.1177/0269758014547993>
- Fohring, S. (2018). What's in a word? Victims on 'victim.' *International Review of Victimology*, 24/2, 151-164.  
<https://journals.sagepub.com/doi/abs/10.1177/0269758018755154?journalCode=irva>

- Fohring, S., & Horsfield, L. (2023). Survivors Speak Out: The Successes and Failures of Hashtag Activism. In V. Canning, G. Martin, S. Tombs, (Ed.), *The Emerald International Handbook of Activist Criminology* (pp. 337-350). Emerald Publishing Limited. <https://doi.org/10.1108/978-1-80262-199-020231022>
- Formula bot. (2025). *Sentiment Analysis Generator*. <https://www.formulabot.com/>
- Foucault, M. (1972). *Archaeology of Knowledge* (A. Sheridan, Trans.). Routledge.
- Fredrickson, B., & Roberts, T. (1997). Objectification Theory: Toward Understanding Women's Lived Experiences and Mental Health Risks. *Psychology of Women Quarterly*, 21(2), 173-206. <https://doi-org.ezproxy.lancs.ac.uk/10.1111/j.1471-6402.1997.tb00108.x>
- Freeden, M. (2017). After the Brexit referendum: revisiting populism as an ideology *Journal of Political Ideologies*, 22(1), 1–11. <https://doi.org/10.1080/13569317.2016.1260813>
- Fricker, M. (2007). *Epistemic injustice power and the ethics of knowing*. Oxford University Press. <https://ebookcentral.proquest.com/lib/lancaster/detail.action?docID=694009>
- Fritze, M. Völckner, F., & Melnyk, V. (2024). Behavioral Labeling: Prompting Consumer Behavior Through Activity Tags. *Journal of marketing*, 88(4), 22-39. <https://web.p.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=0&sid=293b8743-17db-420c-915e-83af6731f142%40redis>
- Fuhrmans, V. (2018). What #MeToo has to do with the Workplace Gender Gap. *Wall Street Journal*. <https://www.proquest.com/docview/2124345296?parentSessionId=im%2FpPpv9TtfwU4OpouDCpAzCdDPDr2Or6RI7Xuh5s%2BI%3D&pq-origsite=primo&accountid=11979&sourcetype=Wire%20Feeds>
- Galer, S. (2020). The accidental invention of the Illuminati conspiracy. *BBC*. <https://www.bbc.com/future/article/20170809-the-accidental-invention-of-the-illuminati-conspiracy>
- Garcia-Moreno, C., Guedes, A., & Knerr, W. (2012). *Understanding and Addressing Violence Against Women*. <https://iris.who.int/server/api/core/bitstreams/21202819-7c42-4f43-8b83-0986c1eb4d9b/content>
- Gill, G., & Rahman-Jones, I. (2020). Me Too founder Tarana Burke: Movement is not over. *BBC*. <https://www.bbc.co.uk/news/newsbeat-53269751>
- Gillings, M., Mautner, G., & Baker, P. (2023). *Corpus-assisted discourse studies* (First edition. ed.). Cambridge University Press. <https://doi.org/10.1017/9781009168144>
- Goffman, E. (1990). *Stigma : notes on the management of spoiled identity*. Penguin.
- Google. (2024). *Google Books Ngram Viewer, Release Notes*. Google. <https://books.google.com/ngrams/info>
- Google. (n.d.). *My Google Activity, Help*. <https://myactivity.google.com/myactivity?pli=1>
- Gorbatch, A. (2016). Different social media platforms: how do they vary and what does it mean to you? *Social Media Monitoring*. <https://awario.com/blog/social-media-platforms/>

- Gould, E., Scheider, J., & Geier, K. (2016). *What is the gender pay gap and is it real?* Economic Policy Institute. <https://www.epi.org/publication/what-is-the-gender-pay-gap-and-is-it-real/>
- Government Equalities Office. (2019). *Gender equality at every stage: a roadmap for change*. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/821889/GEO\\_GEEE\\_Strategy\\_Gender\\_Equality\\_Roadmap\\_Rev\\_1\\_\\_1\\_.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/821889/GEO_GEEE_Strategy_Gender_Equality_Roadmap_Rev_1__1_.pdf)
- Gröndal, H., Jacobsson, M., & Öckinger, V. (2025). Resisting Rape Myths and Doing Feminism in Legal Processes of Rape Cases. *Social politics*. <https://academic.oup.com/sp/article/32/4/1130/8099917?login=false>
- H M Government. (2021). *The end-to-end rape review report on findings and actions*. (CP 437). H M Government. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/1001417/end-to-end-rape-review-report-with-correction-slip.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1001417/end-to-end-rape-review-report-with-correction-slip.pdf)
- Halliday, M., & Matthiessen, C. M. (2014). *Introduction to functional grammar*. Taylor & Francis Group.
- Halpern, D., Valenzuela, S., & Katz, J. (2017). We Face, I Tweet: How Different Social Media Influence Political Participation through Collective and Internal Efficacy. *Journal of Computer-Mediated Communication*, 22(6), 320-336. <https://academic.oup.com/jcmc/article/22/6/320/4675176>
- Harvey, D. (2005). *A brief history of neoliberalism*. Oxford University Press. <https://ebookcentral.proquest.com/lib/lancaster/detail.action?docID=422896>.
- Hawksey, M. (n.d.). TAGS v6.1.9. [https://docs.google.com/spreadsheets/d/1H8IN1Y13MOhgTHc\\_lt7-sGyy9JWf-MeEeKzb7eH7s-M/copy](https://docs.google.com/spreadsheets/d/1H8IN1Y13MOhgTHc_lt7-sGyy9JWf-MeEeKzb7eH7s-M/copy)
- Headley, S., & Rack, S. (2024). Birmingham City Council agrees huge equal pay deal. *BBC* <https://www.bbc.com/news/articles/cdd6y6dnpjro>
- Heffer, C., Tracy, K., Ilie, C., & Sandel, T. (2015). *Jury Trial Discourse*. John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118611463.wbielsi164>
- Hersch, J. (2011). Compensating Differentials for Sexual Harassment. *The American Economic Review*, 101(3), 630-634. <https://www-jstor-org.ezproxy.lancs.ac.uk/stable/pdf/29783819.pdf?refreqid=excelsior%3A4d57189d874654b528b734ad138ad1fb>
- ICC enters era of equal prize money for men and women. (2024). *Sportz Power*. <https://go.gale.com/ps/i.do?p=ITOF&u=unilanc&id=GALE%7CA808905200&v=2.1&it=r>
- Jane, E. (2014). "Your a Ugly, Whorish, Slut": Understanding E-bile. *Feminist media studies*, 14(4), 531-546. <https://www-tandfonline-com.ezproxy.lancs.ac.uk/doi/full/10.1080/14680777.2012.741073>

- Jones, L., Chałupnik, M., Mackenzie, J., & Mullany, L. (2022). 'STFU and start listening to how scared we are': Resisting misogyny on Twitter via #NotAllMen. *Discourse, context & media*, 47, 100596. <https://doi.org/10.1016/j.dcm.2022.100596>
- Kilgarriff, A. (1997). Putting frequencies in the dictionary. *International journal of lexicography*, 10(2), 135-155. <https://academic-oup-com.ezproxy.lancs.ac.uk/book/5162?searchresult=1>
- Kingfisher, C. (2016). Studying Gender and Neoliberalism Transnationally: Implications for Theory and Action. In L. Silverstein & E. Lewin (Eds.), *Mapping Feminist Anthropology in the Twenty-First Century* (pp. 256-275). Rutgers University Press. <https://doi.org/10.36019/9780813574318-014>
- Kossew, S. (2019). Zoë Morrison's Music and Freedom (2016). In M. Gámez Fuentes, S. Núñez Puente, & E. Gómez Nicolau (Eds.), *Re-writing women as victims: from theory to practice* (1st ed., pp. 181-193). Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9781351043601-14/rethinking-victim-zo%C3%AB-morrison-music-freedom-sue-kossew?context=ubx&refId=83af9827-ff7f-4e94-89d6-be395f36b421>
- Kramer, E. (2016). Feminist Linguistics and Linguistic Feminisms. In, & L. Silverstein (Eds.), *Mapping Feminist Anthropology in the Twenty-First Century* (pp. 65). Rutgers University Press. <https://doi.org/10.36019/9780813574318-005>
- Krawcheck, S. (2020). Time's Up for Wall Street. *Marie Claire*. <https://www.marieclaire.com/politics/a30364077/times-up-wall-street/>
- Lakoff, R., & Bucholtz, M. (2004). *Language and woman's place : text and commentaries* (M. Bucholtz, Ed.). Oxford University Press. <https://ebookcentral.proquest.com/lib/lancaster/detail.action?docID=422820>
- Lamphere, L. (2016). Feminist Anthropology Engages Social Movements: Theory, Ethnography, and Activism. In L. Silverstein & E. Lewin (Eds.), *Mapping Feminist Anthropology in the Twenty-First Century*. Rutgers University Press. <https://doi.org/10.36019/9780813574318-014>
- Lanzmann, C. (1985). *Shoah* [Film].
- Law Commission. (2025). *New measures to improve sexual offences cases*. <https://lawcom.gov.uk/news/new-measures-to-improve-sexual-offences-cases/>
- Lawson, C. (2023). *Just like us : digital debates on feminism and fame* (1st ed.). Rutgers University Press. <https://doi.org/10.36019/9781978830943>
- Lerner, M. (1980). *The belief in a just world : a fundamental delusion*. Plenum Press. [https://doi.org/10.1007/978-1-4899-0448-5\\_2](https://doi.org/10.1007/978-1-4899-0448-5_2)
- Lerner, M. (2015). How the Justice Motive Shapes Our Lives. In P. Nathan, R. Cropanzano, & M. Ambrose (Eds.), *The Oxford Handbook of Justice in the Workplace* (pp. 205). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199981410.001.0001>
- Lerner, M. & Miller, D. (1978). Just world research and the attribution process: Looking back and ahead. *Psychological bulletin*, 85(5), 1030-1051.

<https://web.p.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=0&sid=e99dc2a4-8e53-4383-8e0c-db169fba5b4b%40redis>

- Lewis, J. Hamilton, J., & Elmore, J. (2019). Describing the ideal victim: A linguistic analysis of victim descriptions. *Current psychology (New Brunswick, N.J.)*, 40(9), 4324-4332. <https://link-springer-com.ezproxy.lancs.ac.uk/article/10.1007/s12144-019-00347-1>
- Living Wage Foundation. (2024, May). *Almost 2.2 million women in the UK are struggling to make ends meet in low paid jobs*. <https://www.livingwage.org.uk/news/almost-22-million-women-uk-are-struggling-make-ends-meet-low-paid-jobs>
- Loseke, D. (1997). 'The Whole Spirit of Modern Philanthropy': The Construction of the Idea of Charity, 1912-1992. *Social problems (Berkeley, Calif.)*, 44(4), 425-444. <https://doi.org/10.1525/sp.1997.44.4.03x0230v>
- Loseke, D. (2003). *Thinking about Social Problems : An Introduction to Constructionist Perspectives* (2nd ed.). Taylor & Francis Group. <https://ebookcentral.proquest.com/lib/lancaster/detail.action?docID=3410896>
- Mackinnon, C. (1989). *Toward a feminist theory of the state*. Harvard University Press. <https://feminisminnewterms.wordpress.com/wp-content/uploads/2018/08/mackinnon-1989-toward-a-feminist-theory-of-the-state-copie.pdf>
- Mahoney, M. (1995). Victimization or Oppression? Women's Lives, Violence, and Agency. In M. Fineman & R. Mykitiuk (Eds.), *The public nature of private violence : the discovery of domestic abuse*. (1st ed., pp. 59-93). Routledge.
- Maitra, I. (2018). New Words for Old Wrongs. *Episteme*, 15(3), 345-362. <https://www-cambridge-org.ezproxy.lancs.ac.uk/core/journals/episteme/article/new-words-for-old-wrongs/A453EDC07D91B96E7EF652D0B7CCB899>
- Maitra, I., & McGowan, M. (2010). On Silencing, Rape, and Responsibility. *Australasian journal of philosophy*, 88(1), 167-172. <https://doi.org/10.1080/00048400902941331>
- Maitra, M. (2017). Speech and Silencing. In A. Garry, S. Khader, & A. Stone (Eds.), *The Routledge Companion to Feminist Philosophy* (1st ed., pp. 279-291). Taylor & Francis Group. <https://ebookcentral.proquest.com/lib/lancaster/reader.action?docID=5210075&c=RVBVQg&ppg=390>
- Taylor, C., & Marchi, A. (Eds.). (2018). *Corpus approaches to discourse : a critical review*. Routledge. <http://ebookcentral.proquest.com/lib/lancaster/detail.action?docID=5254458>
- Marsh S. (2023). Half of British female gamers experience abuse when playing online. *The Guardian*. <https://www.theguardian.com/games/2023/may/25/half-of-british-female-gamers-experience-abuse-when-playing-online>
- McEney, A., McGlashan, M., & Love, R. (2015). Press and social media reaction to ideologically inspired murder: The case of Lee Rigby. *Discourse & communication*, 9(2), 237-259. <https://doi.org/10.1177/1750481314568545>

- McEnery, A., Xiao, R., & Tono, Y. (2006). *Corpus-Based Language Studies, An advanced resource book*. Routledge.
- McKinsey & Company, & LeanIn.org. (2020). *Women in the Workplace*.  
<https://ebookcentral.proquest.com/lib/lancaster/reader.action?docID=5210075&c=RVBVQg&ppg=390>
- McNaughton Nicholls, C., Mitchell, M., Ian, S., Webster, A., & Hester, M. (2012). *Attitudes to Sentencing Sexual Offences*. S. Council. [https://www.sentencingcouncil.org.uk/wp-content/uploads/Attitudes\\_to\\_Sentencing\\_Sexual\\_Offences\\_web1.pdf](https://www.sentencingcouncil.org.uk/wp-content/uploads/Attitudes_to_Sentencing_Sexual_Offences_web1.pdf)
- Medina, J., & Wood, D. (2008). *Truth: Engagements Across Philosophical Traditions* (J. Medina & D. Wood, Eds. 1. Aufl. ed.). Wiley-Blackwell.  
<https://ebookcentral.proquest.com/lib/lancaster/detail.action?pq-origsite=primo&docID=228605#>
- Merriam-Webster [electronic resource]. (n.d.). Flat-earther. In *merriam-webster.com*.  
<https://www.merriam-webster.com/dictionary/flat-earther>
- Messamore, A., & Paxton, P. (2021). Surviving Victimization: How Service and Advocacy Organizations Describe Traumatic Experiences, 1998–2016. *Social currents*, 8(1), 3–24.  
<https://journals-sagepub-com.ezproxy.lancs.ac.uk/doi/full/10.1177/2329496520948198>
- Milgram, S. (1963). Behavioural Study of Obedience. *Journal of Abnormal and Social Psychology*, 67(4), 371–378.  
[http://www.lphslibrary.org/uploads/7/2/9/6/7296009/milgram\\_obey.pdf](http://www.lphslibrary.org/uploads/7/2/9/6/7296009/milgram_obey.pdf)
- Milli, S., Carroll, M., Wang, Y., Pandey, S., Zhao, S., & Dragan, A. (2025). Engagement, user satisfaction, and the amplification of divisive content on social media. *PNAS nexus*, 4(3), pgaf062. <https://doi.org/10.1093/pnasnexus/pgaf062>
- Ministry of Justice. (2018). *Sexual offending: Ministry of Justice Appendix Tables*. Office of National Statistics.  
<https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/datasets/sexualoffendingministryofjusticeappendixables>
- Moller, A., Sondergaard, H., & Helstrom, L. (2017). Tonic immobility during sexual assault – a common reaction predicting post-traumatic stress disorder and severe depression. *Acta obstetrica et gynecologica Scandinavica*, 96(8), 932.  
<https://doi.org/10.1111/aogs.13174>
- Munro, V. (2008). Of Rights and Rhetoric: Discourses of Degradation and Exploitation in the Context of Sex Trafficking. *Journal of law and society*, 35(2), 240–264. <https://www-jstor-org.ezproxy.lancs.ac.uk/stable/20109793?sid=primo&seq=1>
- Nadone, C. (2018). ‘Women and work’: a cross-linguistic corpus-assisted discourse study in German and in Italian. *Critical Approaches to Discourse Analysis across Disciplines*, 10(1), 167–186. <https://cris.unibo.it/retrieve/e1dcb336-20a4-7715-e053-1705fe0a6cc9/10-Nardone.pdf>

- National Bureau of Economic Research (NBER). (2024, September 2024). *Business Cycle Dating*. <https://www.nber.org/research/business-cycle-dating>
- National Center for Education Statistics. (2021). *FAST FACTS*. National Center for Education Statistics. <https://nces.ed.gov/fastfacts/display.asp?id=98>
- Nexis. (n.d.-a). *Nexis Power Search*. <https://advance.lexis.com/BisNexisPowerSearch/?pdmfid=1519360&crid=277a0317-2356-476a-8ce4-bce2729fef15>
- Nexis. (n.d.-b). *What are Group Duplicates options?* [http://help.lexisnexis.com/tabulaprasa/nexishelp/newsdedupetips-field?lbu=US&locale=en\\_GB&audience=all](http://help.lexisnexis.com/tabulaprasa/nexishelp/newsdedupetips-field?lbu=US&locale=en_GB&audience=all)
- Nexis. (n.d.-c). *What is Nexis?* <https://www.lexisnexis.co.uk/products/nexis.html>
- Nielsen, M., Skov, S., Grundtvig, G., Folker, A., Rugulies, R., Tybjerg A., Madsen, I. (2024). Workplace sexual harassment: a qualitative study of the self-labelling process among employees in Denmark. *International journal of qualitative studies on health and well-being*, 19(1), 2324990-2324990. <https://doi.org/10.1080/17482631.2024.2324990>
- OECD. (2020). *Gender wage gap chart*. OECD Publishing. <https://www.oecd.org/en/data/indicators/gender-wage-gap.html>
- OECD. (2023). *Reporting Gender Pay Gaps in OECD Countries : Guidance for Pay Transparency Implementation, Monitoring and Reform* (1st ed.). OECD Publishing. <https://doi.org/https://doi.org/10.1787/ea13aa68-en>
- Office for National Statistics. (2019). *Gender pay gap in the UK: 2019, Differences in pay between women and men by age, region, full-time and part-time, and occupation. [Statistical Bulletin]*. <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/bulletins/genderpaygapintheuk/2019>
- Office for National Statistics. (2023). *Differences in pay between women and men by age, region, full-time and part-time, and occupation*. <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/bulletins/genderpaygapintheuk/2023>
- Office for National Statistics. (2025a). *Census 2021 gender identity estimates for England and Wales, additional guidance on uncertainty and appropriate use*. <https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/genderidentity/articles/census2021genderidentityestimatesforenglandandwalesadditionalguidanceonuncertaintyandappropriateuse/2025-03-26#glossary>
- Office for National Statistics. (2025b). *Sexual offences victim characteristics, England and Wales: year ending March 2025*. <https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/articles/sexualoffencesvictimcharacteristicsenglandandwales/yearendingmarch2025#sex>
- Office for National Statistics (ONS). (2020). *National life tables – life expectancy in the UK: 2017 to 2019*.

<https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/2017to2019>

- Oxford English Dictionary [electronic resource]. (n.d.-a). Actually. In *oed.com*.  
[https://www.oed.com/dictionary/actually\\_adv?tab=meaning\\_and\\_use#19009750](https://www.oed.com/dictionary/actually_adv?tab=meaning_and_use#19009750)
- Oxford English Dictionary [electronic resource]. (n.d.-b). Already. In *oed.com*.  
[https://www.oed.com/dictionary/already\\_adj?tab=meaning\\_and\\_use#6212081](https://www.oed.com/dictionary/already_adj?tab=meaning_and_use#6212081)
- Oxford English Dictionary [electronic resource]. (n.d.-c). Bandwagon. In *oed.com*.  
[https://www.oed.com/dictionary/bandwagon\\_n?tab=meaning\\_and\\_use#28481579](https://www.oed.com/dictionary/bandwagon_n?tab=meaning_and_use#28481579)
- Oxford English Dictionary [electronic resource]. (n.d.-d). Bitch. In *oed.com*.  
[https://www.oed.com/dictionary/bitch\\_v?tab=meaning\\_and\\_use#28481579](https://www.oed.com/dictionary/bitch_v?tab=meaning_and_use#28481579)
- Oxford English Dictionary [electronic resource]. (n.d.-e). Blame. In *oed.com*.  
[https://www.oed.com/dictionary/blame\\_v?tab=meaning\\_and\\_use#19044533](https://www.oed.com/dictionary/blame_v?tab=meaning_and_use#19044533)
- Oxford English Dictionary [electronic resource]. (n.d.-f). Bloody hell. In *oed.com*.  
[https://www.oed.com/dictionary/hell\\_n?tab=meaning\\_and\\_use-paywall#124378588](https://www.oed.com/dictionary/hell_n?tab=meaning_and_use-paywall#124378588)
- Oxford English Dictionary [electronic resource]. (n.d.-g). Brand. In *oed.com*.  
[https://www.oed.com/dictionary/brand\\_v?tab=meaning\\_and\\_use#15069279](https://www.oed.com/dictionary/brand_v?tab=meaning_and_use#15069279)
- Oxford English Dictionary [electronic resource]. (n.d.-h). Brave. In *oed.com*.  
[https://www.oed.com/dictionary/brave\\_adj?tab=meaning\\_and\\_use#14640324](https://www.oed.com/dictionary/brave_adj?tab=meaning_and_use#14640324)
- Oxford English Dictionary [electronic resource]. (n.d.-i). Bubble. In *oed.com*.  
[https://www.oed.com/dictionary/bubble\\_n?tab=meaning\\_and\\_use#116058426](https://www.oed.com/dictionary/bubble_n?tab=meaning_and_use#116058426)
- Oxford English Dictionary [electronic resource]. (n.d.-j). Certain. In *oed.com*.  
[https://www.oed.com/dictionary/certain\\_adj?tab=meaning\\_and\\_use#9733558](https://www.oed.com/dictionary/certain_adj?tab=meaning_and_use#9733558)
- Oxford English Dictionary [electronic resource]. (n.d.-k). Champion. In *oed.com*.  
[https://www.oed.com/dictionary/champion\\_n?tab=meaning\\_and\\_use#9794333](https://www.oed.com/dictionary/champion_n?tab=meaning_and_use#9794333)
- Oxford English Dictionary [electronic resource]. (n.d.-l). Declare. In *oed.com*.  
[https://www.oed.com/dictionary/declare\\_v?tab=meaning\\_and\\_use#7336000](https://www.oed.com/dictionary/declare_v?tab=meaning_and_use#7336000)
- Oxford English Dictionary [electronic resource]. (n.d.-m). Divide and rule. In *oed.com*.  
[https://www.oed.com/dictionary/divide\\_v?tab=meaning\\_and\\_use](https://www.oed.com/dictionary/divide_v?tab=meaning_and_use)
- Oxford English Dictionary [electronic resource]. (n.d.-n). Enough. In *oed.com*.  
[https://www.oed.com/dictionary/enough\\_adj?tab=meaning\\_and\\_use#5450522](https://www.oed.com/dictionary/enough_adj?tab=meaning_and_use#5450522)
- Oxford English Dictionary [electronic resource]. (n.d.-o). Freak show. In *oed.com*.  
[https://www.oed.com/dictionary/freak-show\\_n?tab=meaning\\_and\\_use#1370962010](https://www.oed.com/dictionary/freak-show_n?tab=meaning_and_use#1370962010)
- Oxford English Dictionary [electronic resource]. (n.d.-p). Glass ceiling. In *oed.com*.  
[https://www.oed.com/dictionary/glass-ceiling\\_n?tab=meaning\\_and\\_use#2973508100](https://www.oed.com/dictionary/glass-ceiling_n?tab=meaning_and_use#2973508100)
- Oxford English Dictionary [electronic resource]. (n.d.-q). Lecture. In *oed.com*.  
[https://www.oed.com/dictionary/lecture\\_v?tab=meaning\\_and\\_use#39605733](https://www.oed.com/dictionary/lecture_v?tab=meaning_and_use#39605733)

- Oxford English Dictionary [electronic resource]. (n.d.-r). Lie. In *oed.com*.  
[https://www.oed.com/dictionary/lie\\_n1?tab=meaning\\_and\\_use#39430865](https://www.oed.com/dictionary/lie_n1?tab=meaning_and_use#39430865)
- Oxford English Dictionary [electronic resource]. (n.d.-s). Miss out. In *oed.com*.  
[https://www.oed.com/dictionary/miss\\_v1?tab=phrasal\\_verbs#36358653](https://www.oed.com/dictionary/miss_v1?tab=phrasal_verbs#36358653)
- Oxford English Dictionary [electronic resource]. (n.d.-t). Missing. In *oed.com*.  
[https://www.oed.com/dictionary/missing\\_adj?tab=meaning\\_and\\_use#36368560](https://www.oed.com/dictionary/missing_adj?tab=meaning_and_use#36368560)
- Oxford English Dictionary [electronic resource]. (n.d.-u). Moan. In *oed.com*.  
[https://www.oed.com/dictionary/moan\\_v?tab=meaning\\_and\\_use#36473113](https://www.oed.com/dictionary/moan_v?tab=meaning_and_use#36473113)
- Oxford English Dictionary [electronic resource]. (n.d.-v). Perpetrator. In *oed.com*.  
[https://www.oed.com/dictionary/perpetrator\\_n?tab=meaning\\_and\\_use#30914911](https://www.oed.com/dictionary/perpetrator_n?tab=meaning_and_use#30914911)
- Oxford English Dictionary [electronic resource]. (n.d.-w). Playing hard to get. In *oed.com*.  
<https://www.oed.com/search/dictionary/?scope=Entries&q=playing+hard+to+get#1180842040>
- Oxford English Dictionary [electronic resource]. (n.d.-x). Please. In *oed.com*.  
[https://www.oed.com/dictionary/please\\_adv?tab=meaning\\_and\\_use#10437051](https://www.oed.com/dictionary/please_adv?tab=meaning_and_use#10437051)
- Oxford English Dictionary [electronic resource]. (n.d.-y). Ploy. In *oed.com*.  
[https://www.oed.com/dictionary/ploy\\_n3?tab=meaning\\_and\\_use#29696033](https://www.oed.com/dictionary/ploy_n3?tab=meaning_and_use#29696033)
- Oxford English Dictionary [electronic resource]. (n.d.-z). Predator. In *oed.com*.  
[https://www.oed.com/dictionary/predator\\_n?tab=meaning\\_and\\_use#28840983](https://www.oed.com/dictionary/predator_n?tab=meaning_and_use#28840983)
- Oxford English Dictionary [electronic resource]. (n.d.-A). Primal. In *oed.com*.  
[https://www.oed.com/dictionary/primal\\_adj?tab=meaning\\_and\\_use#28333008](https://www.oed.com/dictionary/primal_adj?tab=meaning_and_use#28333008)
- Oxford English Dictionary [electronic resource]. (n.d.-B). Reveal. In *oed.com*.  
[https://www.oed.com/dictionary/reveal\\_v?tab=meaning\\_and\\_use#25502573](https://www.oed.com/dictionary/reveal_v?tab=meaning_and_use#25502573)
- Oxford English Dictionary [electronic resource]. (n.d.-C). Ridiculous. In *oed.com*.  
[https://www.oed.com/dictionary/ridiculous\\_adj?tab=meaning\\_and\\_use#25318259](https://www.oed.com/dictionary/ridiculous_adj?tab=meaning_and_use#25318259)
- Oxford English Dictionary [electronic resource]. (n.d.-D). Segregation. In *oed.com*.  
[https://www.oed.com/dictionary/segregation\\_n?tab=meaning\\_and\\_use](https://www.oed.com/dictionary/segregation_n?tab=meaning_and_use)
- Oxford English Dictionary [electronic resource]. (n.d.-E). Shame. In *oed.com*.  
[https://www.oed.com/dictionary/shame\\_n?tab=meaning\\_and\\_use#23159182](https://www.oed.com/dictionary/shame_n?tab=meaning_and_use#23159182)
- Oxford English Dictionary [electronic resource]. (n.d.-F). Shrug. In *oed.com*.  
[https://www.oed.com/dictionary/shrug\\_v?tab=factsheet#22707624](https://www.oed.com/dictionary/shrug_v?tab=factsheet#22707624)
- Oxford English Dictionary [electronic resource]. (n.d.-G). Should. In *oed.com*.  
[https://www.oed.com/dictionary/shall\\_v?tab=meaning\\_and\\_use#23143244](https://www.oed.com/dictionary/shall_v?tab=meaning_and_use#23143244)
- Oxford English Dictionary [electronic resource]. (n.d.-H). Sly. In *oed.com*.  
[https://www.oed.com/dictionary/sly\\_adj?tab=meaning\\_and\\_use#22333550](https://www.oed.com/dictionary/sly_adj?tab=meaning_and_use#22333550)
- Oxford English Dictionary [electronic resource]. (n.d.-I). So-called. In *oed.com*.  
[https://www.oed.com/dictionary/so-called\\_adj?tab=meaning\\_and\\_use#21920199](https://www.oed.com/dictionary/so-called_adj?tab=meaning_and_use#21920199)

- Oxford English Dictionary [electronic resource]. (n.d.-J). Survivor. In *oed.com*.  
[https://www.oed.com/dictionary/survivor\\_n?tab=meaning\\_and\\_use#19448382](https://www.oed.com/dictionary/survivor_n?tab=meaning_and_use#19448382)
- Oxford English Dictionary [electronic resource]. (n.d.-K). Systemic. In *oed.com*.  
[https://www.oed.com/dictionary/systemic\\_adj?tab=meaning\\_and\\_use#19395029](https://www.oed.com/dictionary/systemic_adj?tab=meaning_and_use#19395029)
- Oxford English Dictionary [electronic resource]. (n.d.-L). Take advantage. In *oed.com*.  
[https://www.oed.com/dictionary/advantage\\_n?tab=meaning\\_and\\_use#215132976](https://www.oed.com/dictionary/advantage_n?tab=meaning_and_use#215132976)
- Oxford English Dictionary [electronic resource]. (n.d.-M). Tend. In *oed.com*.  
[https://www.oed.com/dictionary/tend\\_v2?tab=meaning\\_and\\_use#18898652](https://www.oed.com/dictionary/tend_v2?tab=meaning_and_use#18898652)
- Oxford English Dictionary [electronic resource]. (n.d.-N). Vacuum. In *oed.com*.  
[https://www.oed.com/dictionary/vacuum\\_n?tab=meaning\\_and\\_use#16080711](https://www.oed.com/dictionary/vacuum_n?tab=meaning_and_use#16080711)
- Oxford English Dictionary [electronic resource]. (n.d.-O). Victim. In *oed.com*.  
[https://www.oed.com/dictionary/victim\\_n?tab=meaning\\_and\\_use#15540048](https://www.oed.com/dictionary/victim_n?tab=meaning_and_use#15540048)
- Oxford English Dictionary [electronic resource]. (n.d.-P). Zero tolerance. In *oed.com*.  
[https://www.oed.com/dictionary/zero-tolerance\\_n?tab=meaning\\_and\\_use#13679553](https://www.oed.com/dictionary/zero-tolerance_n?tab=meaning_and_use#13679553)
- Oxford Open Economics. (2024). *Re Dimensions of Inequality: The IFS Deaton Review*. Volume 3, Issue Supplement\_1, 2024. <https://doi.org/10.1093/ooec/odad075>
- Palmer-Mehta, V. (2018). The Subversive Power of Survivor Rhetoric: An Innovative Archive of Survivor Discourse in New York Magazine. *Women's Studies in Communication*, 41(2), 159-182. <https://web-b-ebSCOhost-com.ezproxy.lancs.ac.uk/ehost/pdfviewer/pdfviewer?vid=1&sid=85a7217b-6c37-4384-9ca8-4890bf0cab7f%40sessionmgr103>
- Palomino Manjón, P. (2024). Challenging Discourses of Sexual Violence on X: The Linguistic Representation of Victims and Perpetrators in Kavanaugh's Confirmation Process. *Revista alicantina de estudios ingleses (Internet)*(41), 105-127.  
<https://doi.org/10.14198/raei.2024.41.05>
- Papendick, M., Bohner, G., & Windmann, S. (2017). "Passive victim – strong survivor"? Perceived meaning of labels applied to women who were raped. *PLoS ONE*, 12(5).  
<https://doi.org/10.1371/journal.pone.0177550>
- Pardo, M. (2024). Hate Speech We Live By. In S. xu (Ed.), *The Routledge Handbook of Cultural Discourse Studies* (1st ed., pp. 327 - 345). Taylor & Francis Group.  
<https://ebookcentral.proquest.com/lib/lancaster/detail.action?docID=31222155#>
- Parisot, D. (1999). *Galaxy Quest* M. Johnson & C. Newirth; DreamWorks Pictures.  
[https://archive.org/details/title-01\\_20220430](https://archive.org/details/title-01_20220430)
- Partington, A. (2014). Mind the gaps: The role of corpus linguistics in researching absences. *International journal of corpus linguistics*, 19(1), 118-146. <https://www-jbe-platform-com.ezproxy.lancs.ac.uk/docserver/fulltext/ijcl.19.1.05par.pdf?expires=1743698171&id=id&accname=lancaster%2F1&checksum=7B1FAC251775AEC36F29143CDBFFDFFB>

- PayScale. (2020). *Does Pay Transparency Close the Gender Wage Gap? Pay Equity for Men and Women* [White Paper]. <https://www.payscale.com/content/whitepaper/Pay-Transparency-Closing-Gender-Wage-Gap.pdf>
- Pelinka, A. (2013). Right-Wing Populism: Concept and Typology. In R. Wodak , M. KhosraviNik & B. Mral (Ed.). *Right-Wing Populism in Europe: Politics and Discourse* (pp. 3–22). London: Bloomsbury Academic. <http://dx.doi.org/10.5040/9781472544940.ch-001>
- Peplow, G. (2022). Depp v Heard: How has online abuse of Amber Heard become acceptable? *SkyNews*. <https://news.sky.com/story/depp-v-heard-how-has-online-abuse-of-amber-heard-become-acceptable-12624267>
- Peterson, J. (2018). *The Gender Scandal: Part One (Scandinavia) and Part Two (Canada)* . <https://www.jordanbpeterson.com/political-correctness/the-gender-scandal-part-one-scandinavia-and-part-two-canada/>
- Petter, E. (2019, August 15). *How does Twitter manipulate political sentiments?* <https://www.tradefinanceglobal.com/posts/how-does-twitter-manipulate-political-sentiments/>
- Pew Research Center. (n.d.). *About Pew Research Center*. <https://www.pewresearch.org/about/>
- Phipps, A. (2019). The Fight Against Sexual Violence. *Soundings*, 71(71), 62-74. <https://www.proquest.com/docview/2215486771/fulltextPDF/6881058B4F124A8EPQ/1?accountid=11979&sourcetype=Scholarly%20Journals>
- Platt, L. (2024). Gender. *Oxford Open Economics*, 3(Supplement\_1), i291-i293. <https://doi.org/10.1093/ooec/odad081>
- Rafail, P., & Freitas, I. (2019). Grievance Articulation and Community Reactions in the Men’s Rights Movement Online. *Social Media + Society*, 5(2). <https://journals.sagepub.com/doi/10.1177/2056305119841387>
- Rees, S. (2018). A qualitative exploration of the meaning of the term “survivor” to young women living with a history of breast cancer. *European Journal of Cancer Care*, 27(3), e12847. <https://onlinelibrary.wiley.com/doi/abs/10.1111/ecc.12847>
- Rose, K., & McGrory, C. (2020). *UK Social Media User Statistics for 2019*. Rose McGrory Social Media Limited. <http://www.rosemcgrory.co.uk/2019/06/07/uk-social-media-user-statistics-for-2019/>
- Ross, L. (1977). The Intuitive Psychologist And His Shortcomings: Distortions in the Attribution Process. *Advances in experimental social psychology*, 10, 173-220. [https://doi.org/10.1016/S0065-2601\(08\)60357-3](https://doi.org/10.1016/S0065-2601(08)60357-3)
- SAKI. (2021). *Victim or Survivor: Terminology from Investigation Through Prosecution*. <https://www.sakitta.org/toolkit/docs/Victim-or-Survivor-Terminology-from-Investigation-Through-Prosecution.pdf>
- Schäfer, R., & Bildhauer, F. (2013). *Web corpus construction*. Morgan & Claypool. <https://doi.org/10.2200/S00508ED1V01Y201305HLT022>

- Schlote, S. (2023). History of the term 'appeasement': a response to Bailey et al. (2023). *European journal of psychotraumatology*, 14(2), 2183005. <https://www.tandfonline.com/doi/pdf/10.1080/20008066.2023.2183005>
- Scott, L. (2020). *The Double X Economy: The epic potential of empowering women*. Faber & Faber Limited.
- Sehgel, P. (2016). The Forced Heroism of the 'Survivor'. *New York Times*. <https://www.nytimes.com/2016/05/08/magazine/the-forced-heroism-of-the-survivor.html>
- Semino, E. (2018). Language in newspapers. In J. Culpeper, P. Kerswill, R. Wodak, A. McEnery, & F. Katamba (Eds.), *English Language, Description, Variation and Context* (2nd ed., pp. 379-392). Macmillan Education UK.
- Shakespeare, W. (2001). *Henry VI Part Three*. Oxford University Press.
- Shaw, E., Hegewisch, A., & Hess, C. (2018). *Sexual Harassment and Assault at Work: Understanding the Costs*. [https://iwpr.org/wp-content/uploads/2018/10/IWPR-sexual-harassment-brief\\_FINAL.pdf](https://iwpr.org/wp-content/uploads/2018/10/IWPR-sexual-harassment-brief_FINAL.pdf)
- Shoaf, H., Pierce, H., & López, J. (2024). Witches, Victims, and Villains: #MeToo and the Political Polarization of Sexual Violence. *Violence against women*, 30(8), 1910-1933. <https://journals-sagepub-com.ezproxy.lancs.ac.uk/doi/epub/10.1177/10778012241239949>
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
- Sinclair, J., & Carter, R. (2004). *Trust the text : language, corpus and discourse*. Routledge, <https://doi.org/10.4324/9780203594070>
- Singh, V. (2023). The retention problem: Women are going into tech but are also being driven out. *The Conversation : Economy + Business*. <https://theconversation.com/the-retention-problem-women-are-going-into-tech-but-are-also-being-driven-out-200625>
- SketchEngine. (n.d.). *enTenTen: Corpus of the English Web*. <https://www.sketchengine.eu/ententen-english-corpus/>
- Sloan, J., Yoko Furusho, C., Pemberton, A., Bosma, A., Gracia, J., Cohen, C., Cross, C., Bows, H, Heap, V., Fohring, S., Duggan, M. (2018). *Revisiting the 'Ideal Victim' : Developments in Critical Victimology*. Policy Press. <https://doi.org/10.56687/9781447339151>
- Smart, C. (1989). *Feminism and the Power of Law*. Taylor and Francis. <https://ebookcentral.proquest.com/lib/lancaster/reader.action?docID=168805&c=RVBVQg>
- Solnit, R. (2019). *Whose Story Is This*. Granta.
- Statista. (2024). *Leading countries based on number of X (formerly Twitter) users as of April 2024*. Statista Research Department. <https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/>

- Stubbs-Richardson, M., Rader, N., & Cosby, A. (2018). Tweeting rape culture: Examining portrayals of victim blaming in discussions of sexual assault cases on Twitter. *Feminism & Psychology, 28*(1), 90-108. <https://doi.org/10.1177/0959353517715874>
- Sunderland, J. (2004). *Gendered discourses*. Palgrave Macmillan.
- Swain. (2017). Donald Trump's team defends 'alternative facts' after widespread protests. *The Guardian*. <https://www.theguardian.com/us-news/2017/jan/22/donald-trump-kellyanne-conway-inauguration-alternative-facts>
- Syrda, J. (2020). Spousal Relative Income and Male Psychological Distress. *Personality & social psychology bulletin, 46*(6), 976-992. <https://doi.org/10.1177/0146167219883611>
- The BroadBand Commission for Digital Development. (2015). *Cyber Violence Against Women and Girls: A World-wide Wake-up Call*. [https://www.broadbandcommission.org/wp-content/uploads/2021/02/WGGender\\_Executivesummary2015.pdf](https://www.broadbandcommission.org/wp-content/uploads/2021/02/WGGender_Executivesummary2015.pdf)
- The Equal Pay Act 1970*. (1970). <https://www.legislation.gov.uk/ukpga/1970/41/contents>
- The Equality Act 2010*. (2010). <https://www.legislation.gov.uk/ukpga/2010/15/introduction>
- The Equality Act 2010 (Gender Pay Gap Information) Regulations 2017*. (2017). <http://www.legislation.gov.uk/uksi/2017/172/contents/made>
- The Global Institute for Women's Leadership, & Kings College London. (2021). *International Women's Day*. Mori. [https://www.ipsos.com/sites/default/files/ct/publication/documents/2021-03/final\\_report\\_iwd2021.pdf](https://www.ipsos.com/sites/default/files/ct/publication/documents/2021-03/final_report_iwd2021.pdf)
- The Quentin Tarantino Archives. (2003). *Kill Bill*. The Quentin Tarantino Archives. <https://www.tarantino.info/wiki/kill-bill/>
- Tranchese, A. (2023). *From Fritzl to #metoo: Twelve Years of Rape Coverage in the British Press* (1st ed.). Springer International Publishing AG. <https://ebookcentral.proquest.com/lib/lancaster/detail.action?docID=7233616>
- Tranchese, A. (2024). *Reporting on rape: Changing the narrative*. <https://www.endviolenceagainstwomen.org.uk/wp-content/uploads/2024/09/FINAL-RESOURCE-REPORTING-ON-RAPE.pdf>
- Translated by Content Engine, L. L. C. (2023). FIFA endorses equal prize money for women in World Cups. *CE Noticias Financieras*. [Web page deleted]
- Trinch, S. (2003). *Latinas' narratives of domestic abuse : discrepant versions of violence* (1 ed., Vol. 17). John Benjamins Pub. <https://doi.org/10.1075/impact.17>
- Twitter. (n.d.). *Standard search API*. <https://developer.twitter.com/en/docs/twitter-api/v1/tweets/search/api-reference/get-search-tweets>
- U.S. Equal Employment Opportunity Commission. (n.d.). *Chart of Risk Factors for Harassment and Responsive Strategies*. <https://www.eeoc.gov/chart-risk-factors-harassment-and-responsive-strategies>

- U.S. Equal Employment Opportunity Commission. (1964) *Title VII of the Civil Rights Act of 1964*. <https://www.eeoc.gov/statutes/title-vii-civil-rights-act-1964> (eeoc.gov)
- UN-Women (United Nations Entity for Gender Equality and the Empowerment of Women). (2025). *Women's Rights in Review 30 Years After Beijing*. <https://www.unwomen.org/en/digital-library/publications/2025/03/womens-rights-in-review-30-years-after-beijing>
- United Nations Human Rights office of the High Commissioner. (1985) *Declaration of Basic Principles of Justice for Victims of Crime and Abuse of Power*. <https://www.ohchr.org/en/professionalinterest/pages/victimsofcrimeandabuseofpower.aspx>
- Urban Dictionary [electronic resource]. (n.d.-a). Is that all you got? In *urbandictionary.com*. <https://www.urbandictionary.com/define.php?term=that%27s%20all%20you%20got%3F>
- Urban Dictionary [electronic resource]. (n.d.-b). Sitting on a gold mine. In *urbandictionary.com*. <https://www.urbandictionary.com/define.php?term=Sitting%20on%20a%20gold%20mine>
- Urban Dictionary [electronic resource]. (n.d.-c). Pick me girl. In *urbandictionary.com*. <https://www.urbandictionary.com/define.php?term=pick+me+girl>
- US Bureau of Labor Statistics. (2020). *Highlights of women's earnings in 2020* (1094). <https://www.bls.gov/opub/reports/womens-earnings/2020/>
- Van der Tak, E., & Harambam, J. (2024). Reclaiming agency against the new world order: Beyond the coping mechanism frame in conspiracy theory studies. *European journal of cultural and political sociology (Print)*, 11(4), 489-516. <https://www.tandfonline.com/doi/full/10.1080/23254823.2024.2352455#abstract>
- Van Dijk, J. (2009). Free the Victim: A Critique of the Western Conception of Victimhood. *International Review of Victimology*, 16(1), 1-33. <https://journals-sagepub-com.ezproxy.lancs.ac.uk/doi/abs/10.1177/026975800901600101>
- Van Leeuwen, T. (2008). *Discourse and practice new tools for critical discourse analysis* (O. Oxford Scholarship, Ed.). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195323306.003.0006>
- Vandergriff, I. (2013). Emotive communication online: A contextual analysis of computer-mediated communication (CMC) cues. *Journal of pragmatics*, 51(May), 1-12. <https://www-sciencedirect-com.ezproxy.lancs.ac.uk/science/article/pii/S037821661300057X>
- Wagner, D. & Berger, J. (1997). Gender and Interpersonal Task Behaviors: Status Expectation Accounts. *Sociological perspectives*, 40(1), 1-32. <https://doi.org/10.2307/1389491>
- Walker A., & Matt. (2017, 2022). *Mocking SpongeBob*. <https://knowyourmeme.com/memes/mocking-spongebob>

- West, K., Greenland, K., & Laar, C. (2021). Implicit racism, colour blindness, and narrow definitions of discrimination: Why some White people prefer 'All Lives Matter' to 'Black Lives Matter'. *British journal of social psychology*, 60(4), 1136-1153. <https://bpspsychub.onlinelibrary.wiley.com/doi/pdf/10.1111/bjso.12458>
- Westfall, R., Millar, M., & Lovitt, A. (2019). The Influence of Physical Attractiveness on Belief in a Just World. *Psychological reports*, 122(2), 536-549. <https://doi.org/10.1177/0033294118763172>
- Wikipedia. (n.d.-a). *Hanlon's razor*. Wikipedia. [https://en.wikipedia.org/wiki/Hanlon%27s\\_razor](https://en.wikipedia.org/wiki/Hanlon%27s_razor)
- Wikipedia. (n.d.-b). *Sallie Krawcheck*. Wikipedia. [https://en.wikipedia.org/wiki/Sallie\\_Krawcheck](https://en.wikipedia.org/wiki/Sallie_Krawcheck)
- Williamson, J., & Serna, K. (2018). Reconsidering Forced Labels: Outcomes of Sexual Assault Survivors Versus Victims (and Those Who Choose Neither). *Violence Against Women*, 24(6), 668-683. <https://doi-org.ezproxy.lancs.ac.uk/10.1177/1077801217711268>
- Wojcik, S., & Hughes, A. (2019, APRIL 24, 2019). *Sizing Up Twitter Users*. Pew Research Centre. <https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/>
- Wong, W. (2025). Discursive construction of legitimation and delegitimation: a comparative critical discourse analysis of a public debate over Singapore's capital punishment. *Social semiotics*, 1-19. <https://www-tandfonline-com.ezproxy.lancs.ac.uk/doi/full/10.1080/10350330.2025.2543092#d1e96>
- World Economic Forum. (2025). *Global Gender Gap Report 2025*. <https://www.weforum.org/publications/gender-gap-report-2025/>.
- Yapp, E. (2025). Normal Trauma and Abnormal Diagnosis: A Discursive Analysis of Sexual Violence Law and Policy in England and Wales. *Feminist legal studies*. <https://link.springer.com/article/10.1007/s10691-025-09588-x>
- YWCA Regina. (n.d.). *Pyramid of Sexual Violence*. YWCA Regina. <https://ywcaregina.com/programs/education/upstander/>
- Zappavigna, M. (2012a). Ambient Affiliation: A Linguistic Perspective on Twitter. In S. I. R. Methods (Ed.), *SAGE Internet Research Methods* (pp. 193-214). SAGE Publications, Limited. <https://doi.org/10.4135/9781446268513>
- Zappavigna, M. (2012b). *Discourse of Twitter and Social Media: How We Use Language to Create Affiliation on the Web*. Continuum. [https://doi.org/10.1007/978-3-319-17948-3\\_13](https://doi.org/10.1007/978-3-319-17948-3_13)
- Zappavigna, M. (2017). Chp 8. Twitter. In C. Hoffman, & Bublitz, W. (Ed.), *Pragmatics of social media*. De Gruyter. <https://doi.org/10.1515/9783110431070-008>.
- Zhang, K., & Zhuang, H. (2023). Discursive delegitimation of homosexuality on Chinese social media. *Frontiers in psychology*, 14, 1178572. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10521725/>