





*"The maintenance of cohesion should never be taken for granted... The responsibility for fostering cohesion is one shared by us all."*

**Mark Drakeford**

*This page is intentionally left blank.*

## Abstract

---

This thesis investigates the intricate dynamics of community cohesion in Wales and the role played by digital technologies in shaping it. Community cohesion is essential to any functioning modern society, providing the social fabric that binds individuals and groups together. Various emerging challenges have threatened community cohesion in recent years, encompassing cultural polarisation, migration, economic inequality, and numerous others. These factors have created an environment where tensions between groups can escalate, leading to decreased tolerance and subsequent intergroup conflicts and social unrest (Legewie, 2013). While fundamental causal factors and spatial dimensions have traditionally received the most attention, evidence suggests that trigger events can also amplify tensions, resulting in temporal shocks. Social media has surfaced as an additional catalyst for heightening these tensions, enabling the escalation of intergroup conflicts. However, social media also offers the potential for positive outcomes, such as fostering community building and social support. Given the constantly evolving landscape of cohesion in both digital and offline contexts, this PhD research aims to investigate how cohesion is accomplished in Wales. Specifically, the study examines the current mechanisms to prevent tensions and the potential role of new technologies in these processes. This thesis focuses on the relationship between social media and community cohesion, utilising a multi-disciplinary outlook that includes digital communications, tension studies, and computational criminology.

The findings shed light on the multi-agency partnerships in Wales designed to mitigate community tensions and promote community cohesion. First, the Cohesion Delivery Network (CDN) in Wales was explored. It is important to note that given the absence of a pre-established formal multi-agency network, the researcher relied on their positionality within the Hate Crime Criminal Justice Board to define and establish working relationships that existed within the network. The researcher's unique vantage point provided insight into existing collaborative efforts. Consequently, while no formal cohesion delivery network (CDN) exists the researcher's role shed light on the informal network's composition and dynamics.

The cooperation patterns were analysed in a network analysis to identify gaps and challenges in the network and how digital technologies have influenced and will continue to influence how the network operates. Next, a specific case study example of a trigger event for community tensions was explored: the announcement, duration, and closure of the Penally Asylum accommodation in Wales. This helped identify associative factors behind the manifestation and spread of intergroup tensions and the role that positive counter-speech can play in diffusing them following a local trigger event in both offline and digital contexts. In doing so, the study identified the extent to which particular social actors are endorsed or rejected and fundamental temporal and social patterns associated with the creation, propagation, and existence of tensions during trigger events. The concluding section of the thesis explored the third sector's potential to employ social media in fostering digital environments marked by tolerance and cohesion to promote inclusive content and community building. Using Twitter as a public space application, the study found that such organisations harness social media to promote cohesion and collective efficacy using various strategies that ultimately interlink to address threats to community cohesion. Various research methods and tools were used to collect quantitative and qualitative data from offline and digital sources, including survey data and extensive social media data scraped from Twitter.

## Acknowledgements

---

I extend my heartfelt gratitude to the individuals and organisations whose support has been instrumental in the completion of my PhD journey. Firstly, I express my sincere thanks to the research participants who, though remaining anonymous, played a crucial role in the success of this study. Their willingness to engage in the research and share valuable insights has greatly enriched the depth and scope of my work.

To my friends and family, I am profoundly grateful for your unwavering support throughout this process. Your encouragement, understanding, and personal assistance have been the cornerstone of my resilience.

I owe a debt of gratitude to my dedicated supervisors, Matt Williams and Pete Burnap, whose guidance has been invaluable. They not only exposed me to new opportunities in the academic world but also played a pivotal role in my personal and professional development. Their unwavering support in shaping and refining this thesis has been immeasurable. I am also grateful to my annual reviewer, Luca Giommoni, for going above and beyond in his commitment to supporting my academic endeavours. Special thanks to Alyson Lewis, my Associate Fellowship teaching mentor, for invaluable support and guidance beyond the teaching role. Her experience as a former PhD candidate enriched our interactions, providing insights and motivation throughout this journey.

Finally, I extend my thanks to the generous funding and sponsorship from the Welsh Government and The Police. Their financial support not only enabled me to pursue this PhD but also exposed me to real-world practices associated with my work. Being invited to join the Hate Crime and Criminal Justice Board was a testament to the impact of their support, allowing my research to reach a wider audience and engage with diverse stakeholders. I am sincerely thankful for the opportunities and experiences this sponsorship provided.

# Table of Contents

---

<b>Abstract</b> .....	<b>iii</b>
<b>Acknowledgements</b> .....	<b>v</b>
<b>List of Figures</b> .....	<b>viii</b>
<b>List of Tables</b> .....	<b>x</b>

## Chapter 1

### Introduction

1.1. <i>Introduction</i> .....	1
1.2. <i>Research Objectives and Questions</i> .....	4
1.3. <i>Thesis Structure</i> .....	9

## Chapter 2

### Related Literacy and Conceptual Frameworks

2.1. <i>Introduction</i> .....	10
2.2. <i>A New Dawn: The Emergence of Digital Society</i> .....	10
2.3. <i>A Collective Approach to Community Cohesion and Combatting Tensions, Towards Collective Efficacy</i> .....	13
2.4. <i>Digital Communication Strategies</i> .....	39
2.5. <i>Trigger Events for Community Tensions</i> .....	54
2.6. <i>Summary</i> .....	74

## Chapter 3

### Methodology

3.1. <i>Introduction</i> .....	76
3.2. <i>Mixed Methods Approach</i> .....	76
3.3. <i>Empirical Chapter 1</i> .....	81
3.4. <i>Empirical Chapter 2</i> .....	89
3.5. <i>Empirical Chapter 3</i> .....	105
3.6. <i>Ethical Considerations</i> .....	120
3.7. <i>Summary</i> .....	125

## Chapter 4

### Mapping the Cohesion Delivery Network in Wales

4.1. <i>Introduction</i> .....	126
4.2. <i>Hypotheses</i> .....	127



4.3. Results.....	130
4.4. Discussion.....	141
4.5. Summary .....	156

## Chapter 5

Exploring the Nexus of Online and Offline Tensions:  
A Case Study of the Penally Asylum Accommodation

5.1. Introduction .....	159
5.2. Hypotheses.....	160
5.3. Results.....	165
5.4. Discussion.....	186
5.5. Summary .....	207

## Chapter 6

Promoting Cohesion Through Digital Spaces:  
Strategies and Approaches of the Third Sector

6.1. Introduction .....	210
6.2. Hypotheses.....	211
6.3. Data Variables .....	215
6.4. Results.....	216
6.5. Discussion.....	250
6.6. Summary .....	262

## Chapter 7

Discussion

7.1. Introduction .....	264
7.2. External Mechanisms of Cohesion Delivery .....	264
7.3. Informal Mechanisms of Cohesion Delivery.....	272
7.4. The Third Sector: A Hybrid Approach to Collective Efficacy .....	280
7.5. Cohesion Delivery in Digital Society.....	288

## Chapter 8

Summary

8.1. Introduction .....	301
8.2. Contribution Statement.....	302
8.3. Limitations and Future Research.....	309

<b>Bibliography.....</b>	<b>314</b>
--------------------------	------------

<b>Appendices .....</b>	<b>363</b>
-------------------------	------------

<b>Annexes.....</b>	<b>399</b>
---------------------	------------

## List of Figures

---

Figure 1: Summary of Welsh Government 'Framework for Action' objectives and delivery areas .....	26
Figure 2: Arnstein's (1969) Ladder of Citizen Participation .....	32
Figure 3: Cohen's (1972) stages of deviant events based on Williams and Burnap's (2016) trigger event case study (Woolwich terror attack).....	60
Figure 4: Political causes of tensions .....	65
Figure 5: Literature consensus of trigger event typologies .....	67
Figure 6: Methodological steps taken in Chapter Six.....	91
Figure 7: Issue attention cycle, based on Google search data .....	96
Figure 8: Methodological steps taken in Chapter Five .....	106
Figure 9: Coding indicators for broad strategies .....	110
Figure 10: Digital communication sub-strategies .....	112
Figure 11: Analytical phases for Chapter Six.....	118
Figure 12: Decision flow chart for publication of Twitter communications .....	124
Figure 13: CDN cooperation frequency space.....	133
Figure 14: CDN node relationship rank order .....	135
Figure 15: Time series for Penally Tweet traffic and anti-refugee posts .....	167
Figure 16: Time series for Penally anti-refugee posts.....	167
Figure 17: Reply node network for Penally (all replies).....	169
Figure 18: Reply node network for Penally (anti-refugee tweet filter).....	169
Figure 19: Google search trend heat map for "Penally Camps" keyword.....	171
Figure 20: Word cloud for all tweets collected under Penally keyword (top 50) .....	172
Figure 21: Anti-refugee attitudes coefficients by agent .....	177
Figure 22: Positive sentiment tweets (0.7+) (top 40 words, with numbers, special characters and stop words removed) .....	188
Figure 23: Word cloud for anti-refugee attitudes (0.7+) tweets published by far-right political agents .....	194
Figure 24: Word cloud for anti-refugee attitudes (0.7+) tweets published by anti-refugee activists .....	195
Figure 25: Word cloud for all tweets published by pro-refugee activists .....	197
Figure 26: Pro-refugee activist Facebook post .....	199
Figure 27: Tweet strategy prevalence breakdown .....	217
Figure 28: Push strategy attribute radar graph .....	218

Figure 29. News article provided in Twitter link (1) .....	222
Figure 30. News article provided in Twitter link (2) .....	223
Figure 31. Figure 31: News article provided in Twitter link (3) .....	224
Figure 32: Pull strategy attribute radar graph .....	226
Figure 33. Tweet image with event information (1) .....	227
Figure 34: Tweet image with event information (2) .....	228
Figure 35. Tweet image with event information (3) .....	229
Figure 36: Link to organisation's YouTube content .....	232
Figure 37: Networking strategy attribute radar graph .....	234
Figure 38: Transactional strategy attribute radar graph .....	236
Figure 39: Clustered column chart showing regression coefficients for strategy factors (all three models) .....	242
Figure 40: Audience engagement coefficients for 'like' sub-strategy model .....	245
Figure 41: Audience engagement coefficients for 'retweet' sub-strategy model .....	247
Figure 42: Audience engagement coefficients for 'replies' sub-strategy model .....	249
Figure 43: News article provided in Twitter link (4) .....	253
Figure 44: Organisation manifesto provided by Twitter link .....	254
Figure 45: Protest event page provided by Twitter link .....	256

## List of Tables

---

Table 1: Design strategies for analysing social relations with examples.....	92
Table 2: FOI figures for hate crime and incident.....	97
Table 3: Perspective API classifiers cases Google.....	99
Table 4: Pearson correlation matrix for Penally Perspective API scores .....	101
Table 5: Twitter account summary .....	109
Table 6: Emoji typology.....	116
Table 7: Hashtag typology based on Saxton et al. (2015: 161) .....	117
Table 8: Descriptive statistics of CDN respondents (n=59).....	130
Table 9: Factor loadings for cluster group cooperation frequency .....	132
Table 10: BCa bootstrap OLS regression predicting CDN network cooperation .....	140
Table 11: Descriptive Statistics .....	165
Table 12: Agent comparison of modalities.....	174
Table 13: OLS regression predicting anti-refugee attitudes (Penally) .....	175
Table 14: Negative binomial regression, predicting like counts (size model) .....	180
Table 15: Negative binomial regression, predicting retweet counts (size model) .....	182
Table 16: Negative binomial regression, predicting reply counts (size model).....	183
Table 17: Study variables.....	215
Table 18: Negative binomial regression, predicting digital engagements.....	239

# Chapter 1

## Introduction

---

### 1.1. Introduction

The turn of the millennium brought a surge in community tensions worldwide encompassing factors such as globalisation, migration, socio-political polarisation, economic inequalities, and the emergence of a digital society (Fiske, 2002; Thomas, 2011; Edwards et al., 2013). These tensions have threatened cohesion and resulted in heightened demands for formal and informal responses. Formal responses encompass policies and programs implemented by governments and non-governmental organisations to promote cohesion. In contrast, informal responses include community-based initiatives that foster social interaction and understanding. This thesis explores both dynamics in Wales and the role of digital technologies in shaping them. Wales is a compelling case study due to its significant political and economic transformation in recent decades. Its population comprises visibly diverse ethnicities, cultures, languages, gender identities, sexual orientations, and disabilities. However, this diversity has been subject to community tensions, exacerbated by rapid social polarisation, often spearheaded by antecedent trigger events (King and Sutton, 2013). It is crucial to define 'community' within the context of this thesis as encompassing all members of society, whether they coexist with cohesion or not. Additionally, the thesis acknowledges sub-communities, particularly those that are marginalised, which are often recognised as community groups, when organised in more formal capacities.

Between 2019 and 2022, a pivotal period marked by post-Brexit adjustments and the onset of the COVID-19 pandemic, Wales found itself navigating a complex landscape of cultural tensions and digital transformations. In this era of uncertainty, the dynamics of community cohesion emerged as a critical focal point, revealing the intricate interplay between temporal, geographical, political, and community factors. During this timeframe, Wales experienced a unique convergence of socio-political

transitions and global crises. The aftermath of Brexit, coupled with the challenges posed by the COVID-19 pandemic, heightened societal anxieties and underscored the importance of cohesive community responses in the face of adversity. These developments shaped the context within which community cohesion initiatives unfolded, influencing both policy priorities and public discourse. The Welsh Government's framework for action, established in 2017, provided a roadmap for community cohesion governance in Wales. Emphasising prevention, victim support, and detection/monitoring, this framework laid the groundwork for multi-agency collaborations aimed at fostering inclusive and resilient communities. Additionally, the allocation of European Transition Fund resources post-Brexit catalysed the formation of cohesion teams tasked with specifically addressing tensions and promoting social cohesion across the nation in the eight police regions.

Geographically, Wales presents a diverse tapestry of landscapes, each with its own socio-cultural dynamics. Urban centres, characterised by increasing ethnic diversity and digital interconnectedness, stood juxtaposed against culturally homogenous rural enclaves. Central to this study's geographical context was the placement of temporary asylum accommodations, in the Penally former barracks. The Penally asylum accommodation, situated in Pembrokeshire, Wales, served as temporary housing for asylum seekers awaiting the outcome of their immigration claims. Primarily composed of former military barracks, the facility housed individuals and families from various countries who had fled persecution or conflict in their home countries. The majority of residents were asylum seekers, including refugees and migrants seeking protection under international law.

Situated in Pembrokeshire, near Tenby, a region known for its cultural homogeneity, the decision to host asylum seekers in this locale ignited local tensions and galvanised community responses. The Penally case served as a microcosm of broader debates surrounding immigration, identity, and community cohesion, highlighting the need for nuanced approaches grounded in local realities, along with the increasing influence that wider social media can have on more local tensions.

Amidst these temporal, geographical, and political dynamics, communities in Wales grappled with varying perceptions and experiences of cohesion. From urban centres

teeming with diversity to rural hamlets steeped in tradition, each community navigated its unique socio-cultural landscape adjacently to the growing influence of digital society. The Penally asylum accommodation saga epitomised the intersection of these dynamics, serving as a poignant reminder of the challenges and opportunities inherent in building cohesive communities in contemporary Wales.

The impact of social media on community dynamics has also been a subject of debate. On the one hand, some scholars view social media as a polarisation amplifier, enabling like-minded individuals to reinforce each other's intolerant views more efficiently and organised (Sunstein, 2017). On the other hand, digital technologies can enhance community cohesion by enabling communication and interaction across diverse groups, leading to better cross-group understanding and tolerance (Vance et al., 2009). Moreover, these technologies can provide a platform for marginalised communities to express their perspectives, promoting inclusivity and mitigating the risk of social exclusion.

This thesis investigates the intricate dynamics of community cohesion in Wales and the role played by digital technologies in shaping it. The study is guided by conceptual concepts of digital society, collective efficacy, denizenship, nodal governance, the justification–suppression model of the expression and experience of prejudice (JSM) and integrated threat theory (ITT), which serve as lenses for the empirical investigations that constitute the core of this work. By analysing the interplay among these concepts, alongside elements such as digital communication strategies and trigger events, within the specific context of Wales, this thesis endeavours to provide valuable insights into how digital technologies can contribute to fostering community cohesion and promoting social justice in a rapidly changing world.

In the following chapter, the central concepts of this thesis are outlined. The research objectives and questions are also presented, along with a detailed explanation of the overall structure. A mixed methods approach combines survey and social media data to address the research problems and objectives. The pragmatic position adopted acknowledges the complexity of the phenomena under investigation and seeks to generate nuanced and context-specific insights. To provide a

comprehensive understanding of the research, Chapters Two and Three explicate the related literature, core conceptual frameworks, and methodological details that guided the empirical investigations.

The findings of this research contribute to the existing literature in several ways, first, by shedding light on the multifaceted nature of cohesion delivery in a digital society, highlighting the roles played by various social actors, including citizens and non-government organisations (NGOs), in maintaining and promoting community cohesion. Second, it illuminates the impact of digital technologies on cohesion delivery, both positive and negative, and the need to develop nuanced approaches to harness the potential of these technologies while mitigating their negative effects. Finally, this study contributes to the literature on nodal governance and trigger events, providing insights into how collective efficacy can be leveraged to counter divisive forces and promote community resilience in both day-to-day life and the face of crisis. Beyond the academic realm, this thesis holds practical significance as the Welsh Government and Police generously funded it. Consequently, the research findings have been regularly disseminated to these stakeholders and members of the Hate Crime Criminal Justice Board, on which I am a participant. The outcomes of this study have been integrated into a comprehensive network involving a wide array of stakeholders.

## **1.2. Research Objectives and Questions**

This thesis consists of three distinct research chapters that each investigates varying aspects of community cohesion in Wales through the lens of a digital society. The individual research inquiries address different but cognate areas, and the central primary research question drives all three:

“How is community cohesion maintained and promoted in contemporary Wales?”

While the data analysed in each chapter is specific to Wales, the research findings hold significance beyond its boundaries, as practitioners in other parts of the UK can apply the measures and successes in addressing community tensions. Specifically,



the findings about digital spaces, typically not restricted by spatial and temporal dimensions (Castells, 2009), may have a broader application. The following section provides a detailed overview of each research chapter and the associated secondary research questions before presenting a statement about the interconnectedness of the research and its contributions to the primary research question. The secondary research questions are underpinned by a series of hypotheses, expounded upon in their respective chapters. These hypotheses draw on pertinent literature and conceptual frameworks, which are investigated in the next chapter.

### ***1.2.1. Empirical Chapter 1***

*What patterns exist in contemporary Wales's multi-agency network responsible for cohesion delivery?*

The first empirical chapter investigates the cooperation patterns within the Welsh CDN and how these patterns relate to the stakeholders' focus on protected characteristics. The chapter provides an in-depth understanding of how social actors collaborate to maintain community cohesion in Wales. Specifically, the chapter emphasises the significance of a nodal governance multi-agency partnership that includes citizen participation or denizenship in achieving community cohesion.

Furthermore, the chapter examines the impact of the COVID-19 pandemic on the day-to-day operations and cooperation patterns of the CDN and how stakeholders responded to these challenges. Additionally, the chapter analyses the role of social media in community tensions and identifies the requirements for any potential tension monitoring tool that could be implemented to mitigate such tensions. By examining these factors, the chapter aims to contribute to the body of knowledge on community cohesion in Wales and inform policy and practice for achieving effective community cohesion in the age of digital society.

### **1.2.2. Empirical Chapter 2**

*What factors influence public discourse and attitudes towards refugees in response to a localised trigger event, using the example of the Penally asylum accommodation in Wales, and how can collective efficacy from public bodies and citizens be leveraged to and promote inclusive attitudes?*

The second empirical chapter of this thesis investigates the dynamics of community tensions and responses in the aftermath of a localised trigger event - the establishment of Penally asylum accommodation in Wales. The chapter examines how various social actors, including politicians, the media, civil society organisations, and individual citizens, engage in the maintenance of community cohesion in times of crisis. The chapter uses a multi-level framework to explore the interplay of individual attitudes, group norms, and structural factors that shape collective actions and reactions. In particular, the chapter examines how emerging digital technologies, particularly social media, shape the spread of anti-refugee attitudes and facilitate collective responses. By analysing the responses of different social actors; the chapter offers insights into the conditions under which communities can uphold cohesion in the face of trigger events.

### **1.2.3. Empirical Chapter 3**

*How do different communication strategies interlink and contribute to the effectiveness of minority rights NGOs in addressing social injustices, and promoting community cohesion on social media?*

The final empirical chapter investigates how NGOs, dedicated to counteracting tensions and promoting community cohesion leverage social media platforms to address social tensions. Specifically, it delves into these organisations' communication strategies on social media and how they interconnect to form a comprehensive approach to promote cohesion. This study not only investigates the prevalence and utilisation of various approaches but also includes an analysis of engagement metrics on platforms to assess their relative effectiveness, and through

specific case studies, it demonstrates the efficacy of utilising social media to combat tensions and foster day-to-day community cohesion. As such, the secondary research question in this chapter is:

*How do different communication strategies interlink and contribute to the effectiveness of minority rights NGOs in curtailing tensions, addressing social injustices, and promoting community cohesion on social media?*

#### **1.2.4. Research interconnectedness**

This thesis focuses on community cohesion in Wales and investigates the impact of emerging technologies on cohesion and its delivery. Despite the diversity of the research chapters, they share several common themes that link the analysis. Primarily, each chapter focuses on maintaining or promoting cohesion through a nodal governance multi-agency partnership that includes citizen participation or denizenship. This approach is essential to address the complexity and diversity of issues that affect community cohesion in Wales. For example, the first research chapter emphasises the significance of this partnership in maintaining community cohesion, with a specific emphasis on community integration into the network. The second empirical chapter explores various social actors' roles in the aftermath of a community tension-trigger event. The third chapter examines how NGOs employ social media to counter tensions. The chapters are informed by cognate literature and conceptual frameworks and aim to contribute to understanding community cohesion and emerging technologies in Wales.

The second common theme that interlinks the analysis is the digital society lens through which community cohesion and tensions are explored, particularly focusing on how emerging technologies have impacted cohesion and cohesion delivery. The first empirical chapter, which takes a more general exploration of CDN in offline contexts, considers the potential for future tension monitoring and the impact that COVID-19 had on further digital technology integration into day-to-day operations.

The second chapter investigates how social media<sup>1</sup> can impact public discourse and attitudes towards refugees and influence community cohesion after trigger events. In contrast, the third research chapter examines how minority rights NGOs leverage social media to address social injustices and promote community cohesion.

Finally, each research chapter applies the conceptual framework of collective efficacy to understand better how community cohesion is upheld or promoted. This concept is further elaborated upon in the subsequent chapter, but in broad terms, collective efficacy reflects the extent to which community cohesion exists. Achieving and maintaining collective efficacy involves considering both informal factors that pertain to communities and interpersonal relations, as well as external factors that relate to the multi-agency influences on cohesion and responses to tensions. Specifically, collective efficacy is a complex concept that reflects the ability of communities to come together and work collaboratively to address common challenges and threats. It can be influenced by a wide range of factors, including the quality of social ties, the level of social capital, and the effectiveness of governance mechanisms, among other things. Using this framework to examine community cohesion in Wales, this thesis sheds light on the various factors contributing to promoting or eroding community cohesion and how emerging technologies and social actors can impact these processes.

In summary, the diverse research chapters in the thesis ultimately interlink through the common themes of collective efficacy facilitated by the importance of a nodal governance multi-agency partnership that includes citizen participation or denizenship, the exploration of cohesion and tensions through a digital society lens, and the specific examples of how social actors leverage emerging technologies to maintain community cohesion in Wales.

---

<sup>1</sup> For social media analyses, 'Twitter' is commonly referenced throughout this thesis, reflecting the platform's identity during the research, and writing phases. Subsequent to the study's conclusion the platform underwent a name change and is now denoted as 'X'.

### **1.3. Thesis structure**

This thesis is organised into eight distinct chapters, each with its unique focus and contribution. Chapter Two provides an overview of the related literature and core conceptual frameworks that informed the research. To facilitate a nuanced and context-specific analysis, the literature review and conceptual framework are blended in this thesis, with specific case studies employed to examine broader concepts. Chapter Three provides a detailed account of the research methodology employed, beginning with a discussion of the epistemological approach adopted in the thesis. Core ethical considerations are also addressed, along with the designs, data collection, pre-processing, and analysis methods for the three empirical chapters. Chapters Four, Five, and Six focus on the three empirical investigations. Each chapter provides a detailed analysis of the research findings, with Chapter Seven reintegrating the empirical chapters and discussing the broader implications for cohesion delivery. In particular, this chapter places the findings within the context of the related literature and core conceptual frameworks. Chapter eight summarises the thesis, concluding with an overview of knowledge and methodological contributions to the surrounding work, highlighting identified limitations and providing suggestions for future research.

## Chapter 2

### Related Literacy and Conceptual Frameworks

---

#### 2.1. Introduction

The chapter is divided into four parts, each focusing on distinct concepts of this thesis. Part one begins by exploring the conceptual frameworks of information, networks, and digital society to provide an understanding of the digitisation of modern society. This section delves into how these frameworks shape and influence contemporary life. In part two, the notion of community cohesion is examined through the lens of the collective efficacy conceptual framework. The chapter sheds light on the multifaceted nature of cohesion, encompassing both formal and informal capacities within communities. The third part of the chapter investigates current understandings of digital communication strategies. This is achieved by applying Mergel's communication framework, which offers preliminary insights into how organisations utilise digital technologies, such as social media, to accomplish various ends. This exploration enhances our understanding of the role and impact of digital communication strategies in fostering cohesion. Lastly, part four focuses on exploring the temporal associations with community tensions, characterised as trigger events. This section examines real-world case studies and integrates them with the core conceptual framework and other relevant causal underlying theories (i.e. JSM and ITT). Doing so sheds light on how community tensions can be initiated in digital and offline spaces.

#### 2.2. A New Dawn: The Emergence of Digital Society

In the late 20th century, Giddens (1990) coined the term "late modernity" to describe the notable societal and cultural changes. These changes are characterised by globalisation, deindustrialisation, and technological advancements. This thesis focuses specifically on the impact of technological advancements that led to the

digitalisation of social life. In particular, it focuses on two conceptual frameworks, namely, *information society* (Bell, 1976) and *network society* (Castells, 1997) and their eventual integration into one framework referred to as 'digital society' (Powell et al., 2018).

As we move into the late modern era, technology serves a more significant role in shaping our social lives. According to Giddens (1990), technological advancements have significantly changed how we interact and gather information. This trend began in the 1980s and has had a global impact, regardless of a country's size, economic development, or political ideology (Moore, 1997). Webster (2001) noted the proliferation of computer technologies and media services characterised the transition from industrialism to informality. This post-industrial period is commonly referred to as the emergence of the information society, "a society that organises itself around knowledge in the interest of social control and the management of innovation and change" (Bell, 1976b: 576). This shift in the societal organisation is driven by the increasing importance of information and knowledge in the functioning of society. From an economic perspective, this shift can be understood as a move away from an economy based on materials and towards one based on knowledge.

With the advent of information and communication technologies, the production and distribution of knowledge have become more important than physical goods. As a result, there is less emphasis on manufacturing and more on services, particularly in knowledge-based areas. This shift towards an economy based on knowledge is also reflected in the changing nature of work. Bell (1973) highlights that in post-industrialism, "information trumps 'raw muscle, power or energy'" (p. 576). The increasing importance of information and knowledge in the functioning of society has led to a shift in the types of jobs that are in demand. There is a greater need for jobs that involve creating, manipulating, and disseminating information, such as data analysis, information management, and digital marketing. The 1980s witnessed these post-industrial shifts manifest in technological advancements, such as the increasing popularity of computers. The rise of personal computers and the internet has greatly facilitated the creation, manipulation, and dissemination of information, further driving the shift towards an information-based economy.

Moore (1997) identified several key features of the emerging information society, which include how information is used as an economic resource and the changing nature of individual learning, with information serving a more significant role in decision-making at the individual level, including in areas such as consumer decisions, state entitlements, and personal responsibilities. This increased access to information further blurs the boundaries between formal and informal learning (Teresevičienė et al., 2018). Moreover, Bell (1973) identified the rise of the information sector in the economy, including technological infrastructure such as telecommunications and computers, and the emergence of new industries alongside the expansion of existing ones.

Castells (1997) labelled the economic impacts of the shift towards informalism as the emergence of a global informational economy. However, this is not the only implication of digitalisation. Another significant change was the transformation of social structures, a phenomenon first identified by Stein Braten (1981) as the emergence of a "network society." Braten (1981) defined a network society as one powered and organised by networked information and communications technology. While Braten (1981) was early to identify changes in social structures and networks, it was not until the 1990s that sociologists began to take a more comprehensive look at the phenomenon. Castells (1996) built on Braten's work, arguing that networks have become entirely structured around the importance of information and developments in communication technologies.

Today, networks are more efficient than their hierarchal counterparts and are malleable to complex social change. The ease of access to such content and the speed at which it can be shared resulted in an abrupt shift from hierarchal bureaucracies to network structures. As such, communication within network society became decentralised because anyone can access networks autonomously with a powerful minority's consent. Castells (1997) gave an example of this in practice, arguing that indigenous groups such as the Zapatistas in Mexico use networks to interfere with and disrupt traditional power sources (for other examples, see Garnham, 1998; Webster, 2004)



In its broadest sense, digitalisation refers to integrating information communication technology into work, service, and production (Shahidullah, 2018). The restructuring of social life has facilitated this integration according to digital communications and social media technologies (Wildemeersch & Jütte, 2017). The impact of digitalisation is thus evident in all aspects of human life, including the media, economy, social structures, and culture (Orton-Johnson & Prior, 2013; Teresevičienė et al., 2018). Thus, scholars have begun to argue that the relationship between technology and society should be understood as a unified concept rather than as distinct entities. This perspective is supported by the idea that technology has become deeply ingrained in all aspects of contemporary society, shaping, and being shaped by the social, cultural, and economic contexts in which it is used (Lindgren, 2017). As a result, sociological and criminological research has shifted away from more narrow concepts such as network and information society and towards a broader concept encompassing all dimensions of digitalisation, which has been labelled as the "digital society" (Powell et al., 2018).

### **2.3. A Collective Approach to Community Cohesion, Towards Collective Efficacy**

#### ***2.3.1. Defining Cohesion and Tensions and Related Concepts***

Examining the concept's meaning is important before delving into the theoretical and practical aspects of maintaining community cohesion. Generally, cohesion refers to a dynamic process whereby "group members (have an) inclination to forge social bonds, resulting in members sticking together and remaining united" (Carron, 1982, p. 124). This builds on Festinger's (1950, p 7) definition - "forces acting on members to enter and remain in the group". Many have critiqued Festinger's (1950) definition for explaining how cohesion is established or diminished rather than what it truly encompasses (McLeod & Von Treuer, 2013). As a result, Carron's (1982) definition has gained prominence. Two common divisions- 'social' and 'community' cohesion, are often used interchangeably (Lynch, 2001). Despite this, scholars have argued that it is essential to maintain distinctions between each term. For example, Cattle (2018, p. 50) explains that social cohesion pertains to broader socio-economic

factors. In contrast, community cohesion relates to societal fractures along the lines of identifiable communities, such as those defined by faith or ethnicity, as opposed to social class. Given that this thesis centres on concepts of community tensions predicated on group identities (i.e., protected characteristics, see Home Office, 2010), explaining and maintaining these distinctions is important. Moreover, this thesis employs a broader definition of community tensions that considers tensions at a more general level rather than just event-specific ones.

The definition of community tensions, as articulated by a public body, the Welsh Government (2016), is used: "a state of insecurity, uncertainty and disharmony, which has the potential to threaten peace and stability, and which may lead to disorder". The Welsh Government (2016) expand on this, explaining that tensions can extend to "threats, prejudices, experienced or reported events and actions". Tensions, therefore, encompass hate incidents that fall below the threshold of criminal behaviour, such as microaggressions, bullying, implicit bias, and stereotyping (Schweppe and Perry, 2022). Furthermore, numerous cases, both above and below the threshold, go unreported (Nolan & Akiyama, 1999), indicating unknown "dark figures," in victimisation (Ozalp et al., 2020:3). This thesis examines all aspects of hate rather than solely within the confines of legislative frameworks. Therefore, definitions of tensions encompassing all hate forms are deemed most suitable. Scholars have established a correlation between community tensions and a deficiency of community cohesion. As described in the definitions, the deterioration of order and communal harmony is evidence of this relationship. It is posited that a lack of cohesion within a community can result in increased tensions among its members (Williams et al., 2013).

In the context of community cohesion, various attitudes, and behaviors, based on concepts of hate, prejudice, and discrimination, can manifest such as hate crimes, hate incidents, hate speech, and cyberhate. While this thesis primarily focuses on community cohesion and tensions at a broader level, these associated terms hold significant importance and are referenced intermittently throughout. Therefore, it is essential to comprehensively define and delineate each term, providing an opportunity to understand their nuances and implications.

Perhaps the obvious place to start is by understanding the more foundational concepts such as hate and prejudice and how they are interpreted in this thesis. Hate, in a broad sense, refers to hostility and aversion usually deriving from fear, anger, or a sense that injury has occurred (Brudholm, 2016). It can be exhibited in an emotional response against individuals or groups often identified as “different” or a “threat.” This response usually comes from longstanding biases, leading to verbal abuse and physical attacks. The term “hate” lacks a single, clear definition in law or psychology and is often replaced with alternatives like “bias,” “prejudice,” and “hostility” in legislative and academic discourse (Schweppe, 2021). Thereafter, prejudice is a prejudgment or preconceived opinion, often negative, about a group of people or a cultural practice based on stereotypes. It represents a baseless and often negative attitude toward a particular group of people based on race, ethnicity, or religion. Prejudice is a strong and harmful feeling that leads to unkind actions. This system of attitudes and beliefs fosters stereotypes and is deeply integrated into social attitudes, often affecting behaviors and decisions subconsciously (Schweppe, 2021). Discrimination is the unfair treatment of people or groups resulting from prejudice (Wolfe & Copeland, 1994). It is an action that stems from prejudice and disadvantages people based on their membership in a specific group. Discrimination can occur in many settings, such as employment, education, or public service, perpetuating social inequalities and resulting in tangible consequences for those negatively affected (Schweppe, 2021).

One notable distinction is between hate crimes and hate incidents. Hate incidents are subjective acts of bias and hate towards individuals based on their perceived or actual group membership, which may not rise to the level of a criminal act (Glet, 2009). These behaviors include actions perceived as motivated by hostility or prejudice, such as microaggressions—everyday, often unintended but still harmful actions that carry prejudice against minorities. Conversely, hate crimes are criminal offenses against a person or property driven by the offender’s bias against a specific group (Perry, 2015). These crimes are targeted offenses that send an act of aggression and threat to the victim and the group the victim represents. According to Schweppe, a hate crime is a criminal offense with a hate element driven by bias against the victim’s characteristics, which should be an inherent part of the victim’s identity, such as race, religion, or sexual orientation. In this respect, a hate crime can

be classified as a message crime, maintaining certain social and political statuses or hierarchies (Schweppe, 2021). Finally, cyberhate refers to the use of digital platforms to post threatening, harassing, or grossly offensive language targeting individuals or groups based on characteristics like race, religion, or sexual orientation. Defining cyberhate is complex due to cultural and linguistic variations (Burnap & Williams, 2016). Legal scholars, such as Greenawalt (1989), focus on the expressive value of language in classifying hateful speech, considering criteria like provoking violence, deeply wounding targets, causing offense, and degrading social relationships. While UK legislation addresses many aspects of cyberhate, such as the Public Order Act of 1986 and the Protection from Harassment Act 1997, enforcement has historically been inconsistent. The rise of social media has significantly increased the prevalence of cyberhate, particularly affecting young people (Oksanen et al., 2014; Williams & Wall, 2013).

### ***2.3.2. Introduction to Collective Efficacy***

Collective efficacy is "the process of activating or converting social ties among communities to achieve collective goals, such as public order or control of crime" (Sampson, 2010, p. 802). Collective efficacy is an attribute of communities rather than of individuals. As such, collective efficacy depends on the "conditions of mutual trust and solidarity among neighbours" Sampson (2001, p. 95). Social networks, norms, and trust play a significant role in "suppressing norm-deviant behaviour and maintaining social order" (Burton-Smith et al., 2017, p. 608) within a community. It is found that areas with high levels of collective efficacy tend to experience less crime than those with low levels; this extends to many different crime strands, such as adolescent deviance and burglary, violence, and vandalism (Sampson et al., 1997; Sampson, 2012; Higgins and & Hunt, 2016).

Given that the thesis's interest is to explore tensions that manifest socially rather than individually the concept of collective efficacy is pertinent. Evidence suggests that divisive attitudes become an increasingly prevalent and visible problem which threatens cohesion and... collective efficacy" (Ozalp et al., 2020, p. 1). Sampson et al. (1997) argue that collective efficacy is achieved through communities' ability to

"realise common values" by establishing and maintaining social controls. Social controls, defined as "the ability of social groups or institutions to make norms or rules effective" (Reiss, 1951, p. 196), have been a well-studied topic in the social sciences since their inception (Janowitz, 1975). Kornhouser (1978) elaborates on this concept by stating that social controls are based on actual or perceived rewards and penalties, which derive from either conforming to or rejecting societal norms" (Groff, 2015).

In community tension contexts, rewards and penalties can be conceptualised using Crandall and Eshleman's (2003) JSM. The JSM, along with the Integrated Threat Theory (ITT) Stephan (2008), plays a crucial role in determining whether prejudiced feelings that manifest into community tensions are expressed or experienced based on a psycho-social conflict (Williams et al., 2022) between "a desire to express an emotion and, at the same time, maintain values and self-concepts that conflict with prejudice" (Crandall & Eshleman, 2003, p. 414). Suppression forces constitute a complex tapestry of influences, spanning societal norms, personal values, principles, and deeply held beliefs. These subtle forces function as concealed constraints, effectively curbing the public display of antagonistic behaviours that can lead to tensions. Conversely, justification processes act as enablers, providing a conduit for genuine prejudice to find expression. Situational factors, intricate interactions of individual and community experiences, exposure to prevailing ideological currents, and encounters with those with specific attitudes collectively unlock the gates of prejudice, paving the way for its unrestrained public manifestation. Consequently, when justification forces surge while suppression forces recede, the alignment between harbouring prejudiced beliefs and openly exhibiting them intensifies.

Some justification forces revolve around the perception of an outgroup posing a threat to the ingroup, as identified by Crandall and Eshleman (2003). When an outgroup is perceived as a potential threat, it furnishes a rationale for the emergence of prejudice. Two distinct categories of perceived threats exist: tangible threats, encompassing economic concerns, and symbolic threats, derived from cultural differences. These threats influence multiple fronts, affecting both the collective and individual psyche. Members of the ingroup may perceive their vulnerability, their group's susceptibility, or a combination of both in the face of this outgroup threat.

When such a threat is attributed to an outgroup member, it can evoke fear and frustration among ingroup members. These emotional responses, erode the suppression mechanisms while amplifying the justification for expressing prejudice. Importantly, threats need not manifest in overt, observable forms; even the mere perception of a threat, skilfully crafted by public figures or the media, can fan the flames of hatred between groups. It is conceivable, for instance, that the focus on immigration during Trump's election exacerbated the perception of threats, subsequently intensifying the justification for expressing prejudice, which manifested in the further community tensions (Crandal and Eshleman 2003).

While prejudice behaviours can be reduced by limiting internal justifications, scholars note that they are primarily restricted by increasing suppression forces (Crandal & Eshleman, 2003). This conceptual framework will be revisited later in this chapter when exploring trigger events. This section focuses more specifically on suppression forces rather than justification processes. One of the most widely discussed suppression forces is "the public denial of prejudice, through social controls of the expression of prejudice" (Crandal & Eshleman, 2005, p. 245). As such, social control-induced collective efficacy can act as a significant suppressor of divisive behaviours. In this vein, the presence of collective efficacy mirrors the existence of community cohesion. Sampson et al. (1997) suggest that social controls exist in two entities: 'informal mechanisms' and 'formally or externally induced actions' (Sampson et al., 1997, p. 918).

### ***2.3.3. Informal Mechanisms of Collective Efficacy***

Informal mechanisms are divided into two categories: "cohesion and trust" and "informal social control" (Sampson et al., 1997, p. 918). Previous studies have demonstrated that these two control mechanisms are not mutually exclusive (Sampson et al., 1997). Beginning with the former, cohesion and mutual trust were

initially measured by assessing community perceptions of neighbourhoods<sup>2</sup>. A notable addition to these measures is the consideration of "people from different backgrounds getting on well together" (Laurence & Heath, 2008). In this context, scholars have pointed to strong social capital within and between communities as essential to achieving cohesion or collective efficacy (Hawdon and Ryan, 2009). As previously discussed, social cohesion is closely linked to socioeconomic factors such as social class, while community cohesion pertains to the diversity of residents in communities (Cantle, 2018, p. 50). The ability of individuals from diverse backgrounds to coexist harmoniously within a community is vital to promoting collective efficacy from a community cohesion perspective. Much debate has been surrounding the impact of cultural heterogeneity and diversity increases on expressions of prejudice and community cohesion. Scholars later built upon these conflict theories to provide a pessimistic outlook on cultural heterogeneity, arguing that it can negatively impact "attitudes and behaviours, bridging and bonding social capital, public and private connections" (Putnam, 2007, p. 151). Scholars note that these effects can be more profound in rural settings with strong local identity and cultural homogeneity (Garland & Chakraborti, 2006). Heterogeneity encompasses the increasing diversity of other groups. For example, although there is limited research on changes in the proportions of LGBT+ individuals living in the UK, there is ample evidence that visibility increased over the turn of the millennium (Walter, 2003; Burston et al., 2005; Raggins., 2007; Aydin & Ozeren, 2020). As a result, perceived heterogeneity in communities has also increased.

While heterogeneity may negatively impact community cohesion, some studies indicate that these effects can dissipate over time (Twigg et al., 2010). Even the most pessimistic outlooks on cultural heterogeneity acknowledge that negative effects on cohesion are typically most severe in the short term (Putman, 2007). As such, academic interest in the reasons for intergroup conflicts diminishing over time has grown substantially. Scholars now note that areas of high cultural heterogeneity have allowed more investigations (Hewstone, 2015). A growing body of evidence

---

<sup>2</sup> These perceptions were gathered by asking respondents to indicate their agreement with statements such as "the neighbourhood is close-knit," "people in this neighbourhood can be trusted," "people in this neighbourhood generally do not get along with each other," and "people in this neighbourhood share the same values" (Sampson et al., 1997, p. 920).

suggests that regular intergroup contact can help to increase outgroup trust and reduce feelings of perceived threat (McLaren, 2003; Tausch et al., 2007; Pettigrew & Tropp; Dovidio et al., 2017). These findings align with Allport's (1954) intergroup contact theory, which predates conceptualisations of collective efficacy. Allport (1954) proposed that positive interactions between different social groups can promote social change, which he referred to as "positive contact". Such contact can increase knowledge and awareness of other cultures (Dovidio et al., 2017), thereby alleviating perceptions of threats based on inaccurate information (Sternberg & Sternberg, 2008). Chakraborti (2018) points out that social media and digital content can accelerate knowledge campaigns in the digital society. This will be explored further. Cohesion and trust are also established by residents gaining a more accurate understanding of other groups living in their community, allowing them to develop stronger bonds of cohesion and trust that transcend previous feelings of threat and strain (Sampson et al., 1997). These bonds are hinged on integrating sub-communities with 'common norms and values' and 'shared interests' (Ferlander & Timms, 1999; Lynch, 2001; Cattle, 2018). Cohesion and trust can, therefore, be viewed as preventative means of upholding collective efficacy.

In contrast to the proactive process of cohesion and trust, informal control can be perceived as a primarily reactive mechanism for achieving collective efficacy (Warner, 2014). Modern interpretations of informal social control explain that it "consists of the spectrum of actions taken by citizens to signal unacceptable behaviour" (Groff, 2015, p. 90). In terms of collective efficacy, these actions refer to citizen interventions that seek to uphold the common good of their surrounding community (Sampson, 2012). This involves members regulating others to adhere to desired principles. Therefore, informal social controls promote the organic attainment of collective objectives instead of being artificially imposed (Sampson et al., 1997).

Moreover, informal social controls can only flourish in environments of mutual trust. When informal social control was first identified as a proponent of collective efficacy, it was applied to a broader range of issues related to neighbourhood well-being. For example, citizen willingness to intervene was measured by considering five hypothetical scenarios: "children skipping school and hanging out on a street corner, (ii) children spray-painting graffiti on a local building, (iii) children showing disrespect



to an adult, (iv) a fight broke out in front of their house, and (v) the fire station closest to their home was threatened with budget cuts" (Sampson et al., 1997, p. 919).

Examination of these initial indicators for informal social control reveals that it pertains to a wide range of threats to neighbourhood well-being, including anti-social behaviour, violence, and socio-economic decline. Moreover, greater evidence supports claims that community interventions (informal social controls) help alleviate the risks of these factors. Evidence shows that it reduces crime and anti-social behaviours (Bursik & Grasmick, 1993; Sampson et al., 1997; Clear, 2009; Rose & Clea et al., 1998; Petrosino & Pace, 2015; Hoagwood et al., 2018). However, as this thesis specifically explores collective efficacy about community cohesion, i.e., tensions between different sub-demographic groups, the threats to neighbourhood well-being relate more closely to antagonistic expressions. These include cases above a criminal threshold and those below, such as microaggressions, bullying, implicit bias, and stereotyping (Hunt, 2007; Sue et al., 2007; Kopytowska & Baider, 2017; Lehman, 2019).

Informal social controls are divided into two categories: "Direct informal social control," which refers to direct intervention, and "Indirect informal social control," which refers to the mobilisation of formal authorities (Warner, 2007, p. 99). Areas with high cohesion are likely to exhibit both direct and indirect forms of intervention. However, the reasons for choosing one form over the other can vary and may be influenced by factors such as trust in the police, social ties, and individual preferences (Warner, 2007; Tyler & Fagan, 2008). When it comes to addressing discriminatory behaviours, direct informal social controls often involve counter-speech - "citizen-generated counter-speech" is a promising way to promote peaceful, non-polarised discourse (Garland et al., 2022: 1). Levin and McDevitt (2002) argue that informal social responses to community tensions can be beneficial by opposing the specific incident and promoting tolerance within the community. This establishes networks and methods for response, resulting in more efficient and refined future responses. Therefore, informal social controls that address community tensions can serve as preventative and reactive measures, as they signal potential offenders that the community will not tolerate this antagonistic behaviour (Levin & Rabrenovic, 2004). These controls help establish communal norms and values that indicate

intolerant behaviours are unacceptable (Ferlander & Timms, 1999). Ultimately, this strengthens cohesion among community members. Recent research has further explored community responses. Evidence suggests that grassroots, "bottom-up" responses can most effectively mobilise community changes in perceptions of antagonistic behaviours. Reny and Newman (2021) refer to this as the "opinion mobilising effect." The effectiveness of community-driven responses has been well-documented, with communal condemnations having the ability to shape behaviours, values, and norms across various settings, including micro (e.g., libraries), local (e.g., neighbourhoods), regional (e.g., counties and states), national, and even international contexts (Defeis, 1992; Walters, 2014; Chae, 2022).

#### ***2.3.4. Externally Actions of Collective Efficacy***

Traditionally, external influences (see Sampson, 2010) were centralised to public sector nodes (Kickert et al., 1997). However, the turn of the millennium has shifted towards a more pluralised approach involving a wider range of social actors (Crawford, 2002). The origins of this shift can be traced back to the 1980s, during the new public management reforms, which aimed to increase efficiency *via* implementing management models similar to those used in the private sector (Hood, 1991). A key aspect of these reforms was the creation of "single purpose" public sector agencies (Christensen & Lægreid, 2006, p. 4), with specialised roles to enhance the efficiency of public sector operations and emulate the benefits of division of labour in private contexts (Powell, 2000). This led to many new sub-agencies and program offices in Western public sector systems, such as Australia and the UK (Christensen & Lægreid, 2006).

This diversification of public sector actors was referred to as "multi-lateralisation" by Bayley (2001). While the diversification of public sector actors was generally viewed positively, scholars criticised the new public sector reforms for being state-centric and failing to consider broader social implications (Shearing & Wood, 2003). For example, in criminal justice contexts, state-centric approaches were found to prioritise the state's ability to use coercion to ensure compliance rather than addressing the underlying causes of social issues (Shearing & Wood, 2003, p. 402).

In response to these criticisms, a “preventative turn” emerged, which built upon multi-lateralization and shifted towards a system of nodal governance (Crawford, 1997; Garland, 2001; Shearing, 2001; Levi & Williams, 2013). Nodal governance refers to a system in which various actors, including those from the voluntary, private, and community sectors, work in coordination to address issues in social systems (Burriss, 2004). This approach encourages a “diffusion of responsibilities” among multi-agency networks of social actors who previously did not have “regulatory mandates” (Levi & Williams, 2013, p. 422). This pluralisation of systems and inclusion of non-governmental actors was often an organic process but was formalised in UK law through the Crime and Disorder Act in 1998 (Shearing & Wood, 2003).

Despite some critiques of nodal governance (Sedgwick et al., 2020), academics and practitioners have a widely accepted consensus that a more pluralistic approach leads to improved outcomes. For instance, multi-agency networks in crime prevention have demonstrated a reduction in crime rates (Choi & Choi, 2012). To optimise multi-agency partnerships, it is essential to enhance the frequency and effectiveness of cooperation among all stakeholders involved, including those at the policy, community, and enforcement levels (Sedgwick et al., 2020). Pluralistic approaches allow for a more comprehensive understanding of issues within social systems (Burriss, 2004), surpassing the limited perspectives gained by relying solely on the public sector's traditional reach (Crawford, 2002, p. 31). Nodal governance systems build upon the benefits first realised during the new public management reforms of multi-lateralisation, in which a diverse network of partners can specialise in different areas and bring unique knowledge and capacities to the table (Johnston & Shearing, 2003). Specialised knowledge between different nodes reflects the concept of cultural capital (Dupont, 2004).

In addition to cultural capital, Dupont (2004) identified four other forms of capital essential in any nodal network. He argued that although specialist knowledge is crucial, it loses value if not shared among other nodes. Dupont (2004) argued that social capital must also be achieved to ensure knowledge is shared. In multi-agency contexts, social capital refers to how nodes communicate and cooperate (Halpern, 2005). Nodes with high social capital can foster and sustain mutually beneficial

relations with other nodes (Levi & Williams, 2013, p. 422). In this way, social capital can be seen as a means for cultural capital to take effect. Communicating and sharing specialist information allows for a network-wide critical awareness of the social issue (Kagan & Duggan, 2011). However, the willingness of different nodes to share information can be hindered in several ways.

For example, regarding hierarchical issues, social capital can be hindered when organisations at the state level do not share information or cooperate with those at lower levels. One key example is when communities are not effectively engaged or consulted by public sector nodes. In some cases, communities may be entirely neglected; in others, their involvement may take on a more tokenistic form.

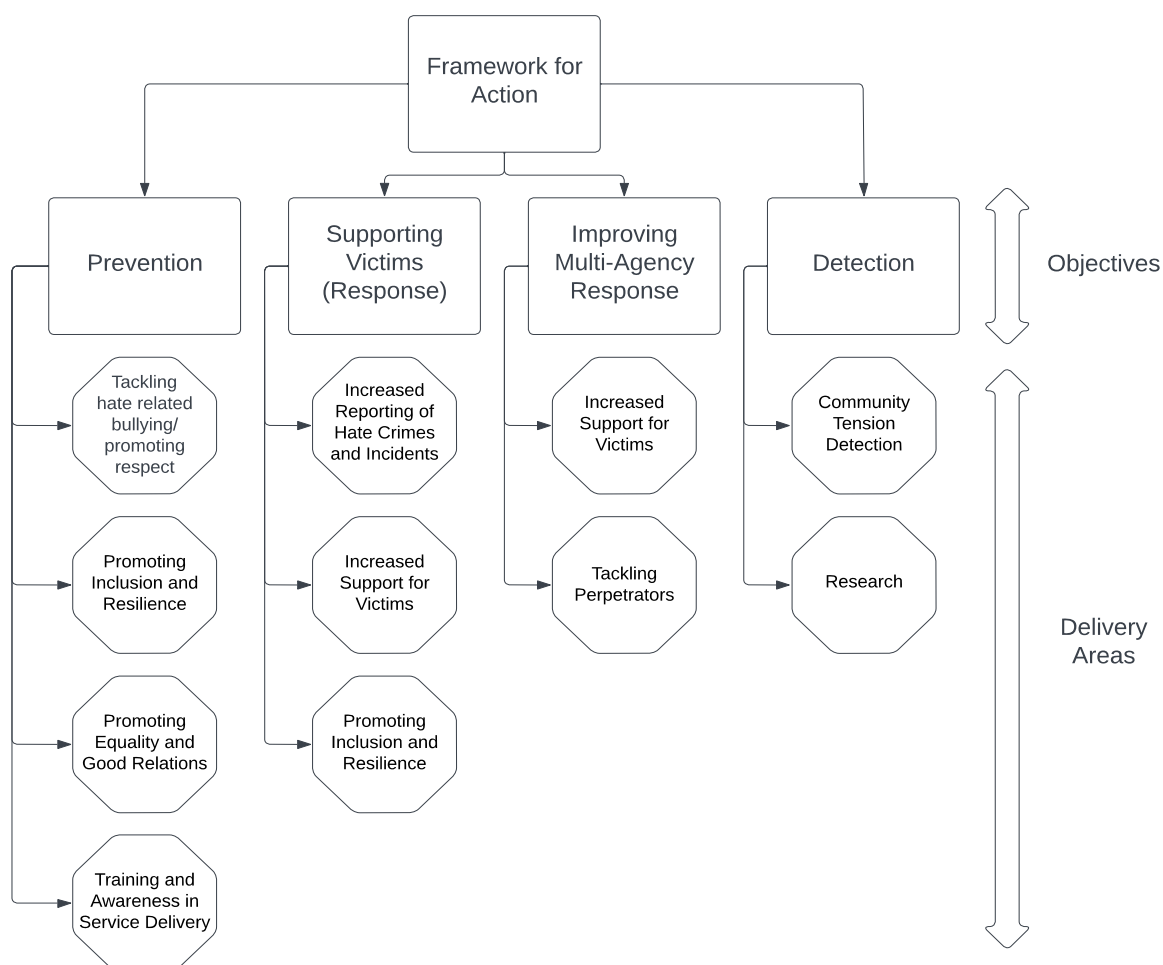
Hierarchical faults in information-sharing do not only occur in the case of community groups. Evidence also suggests that a lack of clarity from program coordinators, such as government bodies, can disconnect between intended projects and their ultimate outcomes (Cheminais, 2009; Edwards & Hughes, 2009). Horizontal information-sharing faults can occur between nodes operating in the same sectors, such as between the police and government (Cheminais, 2009; Thomas, 2010; Levi & Williams, 2013). In this context, non-reciprocal communication refers to instances where one party shares information while the other fails to reciprocate. This can lead to several collaboration-related issues (see Cheminais, 2009; Levi & Williams, 2013).

The idea of 'collective efficacy' is complex, encompassing social and economic issues such as social order and crime control (Weisburd et al., 2015; Sampson, 2017). The level of nodal governance or external involvement required varies depending on the issue. For instance, tackling neighbourhood graffiti may only require the involvement of partners such as police and policymakers (Higgins & Hunt, 2016). However, addressing the multifaceted issue of community tensions, which is the focus of this thesis, may require a more comprehensive range of partners. Many advocate for a multi-agency approach involving "developing partnerships across different spheres of work" to detect, respond to, and prevent, attitudes that may exacerbate community tensions (Chakraborti, 2018: 400).

Community cohesion is achieved when these tensions are minimised and subsided. At the time of writing, academic literature that directly explores multi-agency

networks responsible for cohesion delivery at a national level is scarce. In light of this, the work of Blake et al. (2008) provides an overview of key partners by considering regional case studies. These partners include the public and voluntary sectors (charities and faith groups) and communities. Public agents were further categorised into local councils, the police, and cohesion teams (Blake et al., 2008). These are later employed to analyse the Welsh CDN in Chapter Four. While the literature on how these partners uphold collective efficacy through a cohesion lens is scarce, more has been discussed on how they interact with collective efficacy at a more general level. All of these partners, except community groups, operate by carrying out "externally induced actions" (Sampson et al., 1997, p. 918) that influence collective efficacy. As previously mentioned, externally induced actions manifest in "direct control interventions" and "informal mechanism facilitations". Direct control interventions refer to actions that directly address threats to social order. These tend to take place amongst public sector nodes. Perhaps most notable is by policy practitioners, who contribute to creating laws that actively restrict undesirable behaviours.

Policy practitioners considered in this thesis are the Welsh Government. Although they may not devise many hate legislations, as the UK's central government enacts many, they contribute to cohesion delivery using political capital in many other ways. Political capital enables nodes to influence and create policy, determine public expenditure, or strategise multi-agency responses (Levi & Williams, 2013). The Welsh Government's framework for action, entitled "Tackling Hate Crimes and Incidents" (2017), serves as a prime example of the processes involved in nodal governance and the integration of pre-existing agents into multi-agency networks. The framework outlines four central objectives, each with various sub-delivery areas, demonstrating the government's support of a nodal governance system to improve the multi-agency response to tensions and associated hate crimes and incidents.



*Figure 1: Summary of Welsh Government 'Framework for Action' objectives and delivery areas*

Government actors are integral for coordinating and integrating agents into networks- “one of the keys to generating social goods, and collective efficacy, is institutions that are viewed as legitimate and supported by strong government” (Sampson, 2002, p. 221). According to Posner (2009), one significant agent governments often integrate into these networks is academics, as they possess a significant degree of cultural capital or specialist knowledge, which the government can diffuse to the network through their role as intermediate bodies. This is beneficial because their evidence-based inputs can hold value to the network, yet they do not need to interact with all partners to diffuse them (Kagan & Dugan, 2011). That being said, many academics dispute the current extent of academic integration into such partnerships, describing them as “tick box” and “tokenistic”, resulting in a lack of empiricism guiding the wider network's actions (Chakraborti, 2016: 1)

Moreover, the government holds power to create new specialist roles, as exemplified by the use of the EU transition fund, allocated to support cohesion-related operations in response to the challenges presented by Brexit. This funding has been utilised to establish new roles, such as coordinators and officers, in the pursuit of promoting cohesion on a national level. The framework for action centres around four main objectives: prevention, detection, victim support, and a multi-agency approach. This highlights Wales's multi-lateralisation system of cohesion delivery, creating single-purpose public sector actors (Bayley, 2001; Christensen & Lægneid, 2006).

This approach has temporal and sectoral implications and leads to a more widespread presence of the state across various coordinated frameworks and partners (Zedner, 2007). However, as Kean and Hamilton (2004) noted, there is a potential risk in adopting this approach, as the objectives of different agencies may conflict with the overarching objectives set by the state. This risk has been observed in other British multi-agency partnerships, such as the Preventing Violent Extremism (PVE) network (Thomas, 2010). The network expansion has led to power conflicts driven by competing objectives and turf expansion between local authorities and the police. In conclusion, while multi-agency networks can bring greater cultural capital to the network, they lose any value if social capital, or the ability to share information effectively, is not upheld. In light of this, Chapter Four's main focus of analysis is to understand cooperation trends between key partners, ascertain what kinds of cultural capital are held, and determine whether they are effectively diffused into the network through good social capital.

Government bodies play a crucial role in enacting direct control interventions, such as policies and coordinating networks tasked with specific agendas. They can also engage in informal mechanism facilitations when external agents facilitate mutual trust development and informal social control. These actions are designed to influence community cohesion rather than enforce it. Regarding cohesion, governing bodies achieve this through programs and initiatives to promote diversity, tolerance, and inclusion. These programs serve as "myth-busting exercises" (Blake et al., 2008, p. 74) that expose people to more realistic information about outgroup cultures and ways of life. This increased knowledge and awareness of outgroups (Dovidio et al., 2017) can help reduce falsely predicated feelings of threat previously felt by the

ingroup (Sternberg & Sternberg, 2008). By promoting understanding and acceptance of diversity, governing bodies can foster a sense of cohesion within communities (Gronholm et al., 2017).

The literature on the role of the police in maintaining collective efficacy is extensive, with a significant body of research exploring the actions taken by the police as an external agent to uphold collective efficacy (Hart & Colavito, 2011; Sargeant et al., 2013; Nix et al., 2015; Wiesburd et al., 2015). In the context of community cohesion, this refers to the police enforcing hate crime laws and diffusing tensions. Initially, discussions of the police's ability to build collective efficacy focused on legitimacy as a pivotal factor in their success (LcFree, 1998). Legitimacy in this context is based on the idea that agencies such as the police have an obligation to intervene beyond their self-interest (Kochel, 2012). The police play an extremely important role in maintaining collective efficacy, and high degrees of social capital within communities are integral to their success (Dupont, 2004). The literature suggests that the tactics utilised by law enforcement significantly impact community perceptions of their legitimacy, trustworthiness, and effectiveness and can ultimately either foster or hinder collective efficacy (Scott, 2002).

Notions of trustworthiness in policing have evolved from prior conceptualisations of legitimacy to encompass not only the sense of obligation to obey due to factors such as government endorsement but also the degree to which the police can be trusted (Tyler, 2006). A crucial aspect of police legitimacy is the perceived level of procedural justice in police operations (Sunshine & Tyler, 2003). When trust and legitimacy are established within the police force, they can uphold symbolic capital and improve community relationships (Dupont, 2004). However, traditional policing strategies such as police raids, crackdowns, and minority targeting or profiling can decrease trust and lead to conflict between the police and communities (Gau & Brunson, 2010; Weitzer, 2003; Slocum et al., 2010; Sargeant et al., 2013). Notably, breakdowns of minority group trust in the police has been observed in rural communities (Chakraborti & Garland, 2003), which reflects a large portion of Wales (Scott, 2020). While some scholars argue that the inherent nature of policing will always involve a level of threat and, therefore, limit the relationship between the



police and communities to some extent (DeLeon-Granados, 1999), there is a consensus that certain strategies can be employed to mitigate these.

Police strategies to enhance trust include community policing (Fielding, 1995), problem-oriented policing (Spelman, 1987), intelligence-led policing (Ratcliffe, 2003), and police image work (Mawby, 2002). The former three strategies are directly related to policing practices and have been proposed to increase levels of collective efficacy by promoting self-help and facilitating access to police resources (Velez, 2001). Police image work does not refer to actual processes carried out by the police but rather how they present their operations to the public. These include “activities in which police forces engage and which construct and project images and meanings of policing” (Mawby, 2002, p. 5). Evidence suggests that when relationships are improved between minority groups and the police, hate crimes and incidents are more likely to be reported (Sullaway, 2004). Therefore, it is crucial to establish trust between the police and communities to improve social capital (Kochel, 2012). Moreover, multi-agency partnerships have investigated alternative methods to enhance reporting. One promising approach involves incorporating third-party organisations, such as NGOs, to collect hate crime data (Schweppe et al., 2020).

The final determining factor of the police's influence on collective efficacy is the perceived effectiveness of their actions by the communities they serve (Sargeant et al., 2013). Research suggests that when the police are perceived as effective<sup>3</sup>, communities begin to feel that they can depend on the police more (Silver & Miller, 2004; Warner, 2007; Sargeant et al., 2013). Increased perceptions of police effectiveness can increase engagement in informal social controls, as residents are more likely to intervene or call the police directly (Drakulich et al., 2012; Warner, 2007, p. 99). However, it is important to note that actual effectiveness can differ from perceived effectiveness. In some cases, communities may perceive the police as less effective than they are, a phenomenon called the "reassurance gap" (Millie and Herrington, 2005). But overall, upholding an image of police effectiveness is likewise a means of achieving symbolic capital (Dupont, 2004).

---

<sup>3</sup> i.e., controlling crime and providing timely responses to service calls.

As previously stated, nodal networks comprise diverse stakeholders from various sectors, widely acknowledged as a positive development as it enhances knowledge and capacities, or cultural capital, among the multi-agency response (Dupont, 2004). In the context of collective efficacy, many stakeholders are labelled "external" forces (Bunar, 2007). Crawford (2002) notes that while the network primarily comprises external actors, it also includes informal community groups. The role and significance of communities within networks can vary depending on the network and the social issue being addressed.

In this thesis, which specifically explores cohesion delivery networks, scholars highlight the importance of communities and civil actors playing a significant role (Lepineux, 2005). While the benefits of external actions on community cohesion are evident, scholars have warned that harmful outcomes can result if communities become overly reliant on them (Wells et al., 2006, p. 544). Therefore, a careful balance of formal and informal contributions is advocated in cohesion delivery networks. A strong emphasis should be placed on building vertical relationships between communities and external agents (Blake et al., 2008, p. 12), as the desired outcome of community cohesion and collective efficacy occurs primarily in informal civil society spaces (Bayley & French, 2008). It is therefore deemed crucial to meaningfully engage communities within the network by establishing connections between formal public institutions and informal communities and individuals (Sargeant et al., 2013). This ultimately supports citizens and private actors to govern their spaces, promoting community cohesion and achieving collective efficacy (Garland, 2001).

### ***2.3.5. Community Participation with External Forces***

A useful way of understanding this space is through Arnstein's 'Ladder of Engagement' (1969), which helps locate the level of citizen participation, or denizenship, within multi-agency networks. Arnstein (1969) critiques the tendency of governmental bodies to use vague terms such as "community partnership" that give the appearance of collaboration without actual citizen involvement. He posits that participation exists on a spectrum, with many forms operating well below meaningful

implementation. The ladder comprises eight distinct rungs, which can be grouped into three broader categories of citizen participation: non-participation, degree of tokenism, and degree of citizen power. The lowest category of participation is non-participation, which includes the rungs of "manipulation" and "therapy" (Arnstein, 1969, p. 217). These rungs involve no citizen influence and limited participation. They are regarded as one-way and lack genuine two-way dialogue with communities (Bayley & French, 2008, p. 199). Manipulation refers to the manipulation of citizens by government officials, where citizens are used for the benefit of officials without their input or consent (Arnstein, 1969).

Conversely, therapy refers to providing psychological or social services to citizens to improve their well-being without their participation or input (Arnstein, 1969). The second category is the degree of tokenism, which includes the rungs of "informing," "consultation," and "placation" (Arnstein, 1969, p. 217). Informing refers to providing citizens with information about decisions and plans without allowing them to provide input or feedback (Arnstein, 1969). Consultation refers to seeking citizens' input and feedback on decisions and plans without giving them any real decision-making power (Arnstein, 1969). Placation refers to providing citizens with a sense of involvement or participation without any real impact on the final decision or outcome (Arnstein, 1969). As such, it is the "highest level tokenism because the ground rules allow (citizens) to advise but retain for the powerholders the continued right to decide" (Arnstein, 1969, p. 217). Finally, these rungs are tokenistic because they do not hold any meaningful influence or power. The literature suggests that tokenistic participation, characterised by minimal influence and power, can lead to disillusionment among citizens, resulting in significant reductions in symbolic capital (Brodie et al., 2009). This, in turn, can lead to a cyclical effect of diminishing citizen participation over time.

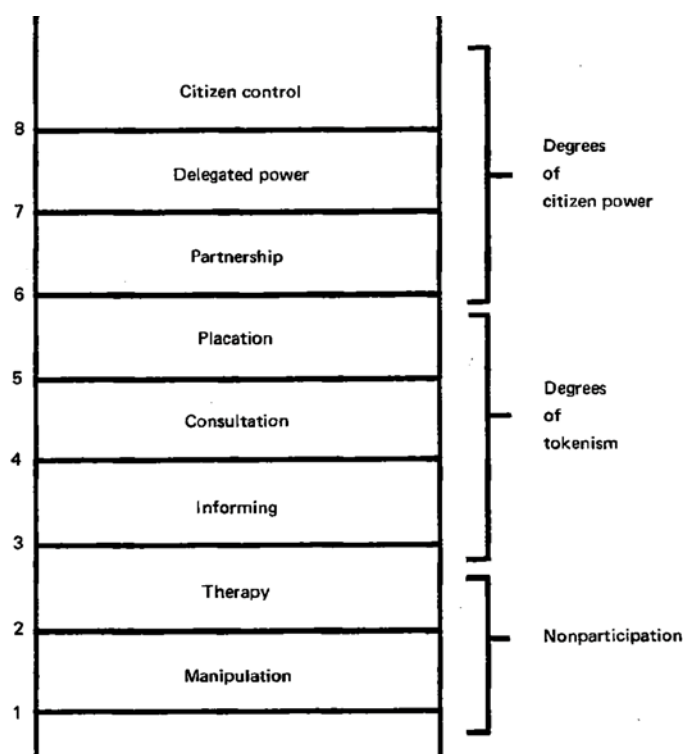


Figure 2: Arnstein's (1969) Ladder of Citizen Participation

The final category is the degree of citizen power, which includes the rungs of "partnership," "delegated power," and "citizen control" (Arnstein, 1969, p. 217). These rungs represent increasing levels of citizenship (Shearing & Wood, 2003). Partnership refers to collaboration and shared decision-making between government officials and citizens. Delegated Power refers to giving citizens decision-making power in specific areas or issues. Citizen Control refers to giving citizens full control over decision-making and governance.

Arnstein (1969) acknowledges that complete citizen control may present its issues but argues that meaningful engagement above a tokenistic threshold is beneficial, especially those focusing on community-based issues. Scholars note that community participation in multi-agency networks can occur at various levels, with certain partners being more suitable depending on the specific network. Legg (2021) conceptualises areas of access to multi-agency networks as "pressure points". These can be understood as network nodes with strong social and symbolic capital in communities (Dupont, 2004). This is characterised by trustworthiness, legitimacy, and a willingness to interact with communities (Scott, 2002). Therefore, the analysis

in Chapter Four examines communities' diverse relationships with external stakeholders within the Welsh CDN to identify potential "pressure points" for future consideration to facilitate community access to the network and promoting more meaningful participation in cohesion practices. According to scholars, meaningful community participation is essential for multi-agency partnerships that focus on cohesion. This is because such participation helps to foster empathy for victims and creates a sense of tolerance and cohesion within the community. Therefore, it is important to encourage active involvement from community members to better understand the issues and work towards a more inclusive and just society (Chakraborti, 2018; Garland et al., 2022).

As noted, many external partners within multi-agency networks are within the public sector. They can contribute significantly to overall delivery in various ways. For instance, many public sector partners possess significant capabilities to influence or even determine policy, which Dupont (2004) calls "political capital." Moreover, creating single-purpose agencies in the public sector has led to greater specialisation of knowledge concerning specific facets of social issues, a concept Dupont (2004) calls "cultural capital." In terms of collective efficacy, it has been emphasised that engaging communities is of paramount importance in promoting community cohesion (Lepineux, 2005). Therefore, fostering social capital between external partners and communities is crucial.

However, many public sector partners have been limited in their ability to achieve high levels of social capital with community groups. This is partly due to cases of public agents unwilling to include communities in operations (Arnstein, 1969) and low levels of symbolic capital (how effective and trustworthy they appear to community groups). One potential "pressure point" (Legg, 2021) for community groups to enter cohesion delivery networks may exist outside of the public sector within the third sector. Scholars have observed that NGOs possess many of the same characteristics and organisational structures as those in the public sector (Salamon & Anheier, 1997; Kendall & Knapp, 1997). This enhances their symbolic capital as professional appearances allow them to establish an image of legitimacy. Furthermore, the third sector also allows for the specialisation of roles or "interest groups" within different agencies (Kelly, 2007), which leads to an accumulation of

cultural capital similar to that found in single-purpose agencies within the public sector (Dupont, 2004).

Although many similarities have been identified with the public sector, scholars have argued that NGOs are more flexible and, therefore, more adaptable to change due to their "structural and ambiguous organisational characteristics" (Kelly, 2007, p. 1010). As a result, third sector partners may possess a "comparative advantage" over other partners within multi-agency networks to effectively engage community groups in meaningful ways (Billis & Glennester, 1998, p. 79). By adopting a more pragmatic approach, the third sector can strengthen social capital with community groups as they present more relatable and normative environments with "communitarian" undertones that are more appealing for community groups to interact with (Percy-Smith, 2003). Furthermore, it has been argued that the degree of flexibility, coupled with collectivist approaches, leaves the third sector "uniquely placed to manufacture and sustain self-supporting network ties" with communities (Kelly, 2007, p. 1010).

Scholars argue that the role of the third sector in fostering collective efficacy extends beyond its ability to bridge communities and external partners. The third sector is uniquely positioned as a hybrid of external institutions and communities (Kelly, 2007, p. 1010). In this vein, they are presented with an unprecedented opportunity to promote networking "with and between communities" (Blake et al., 2008, p. 74). As such, they not only act as a pressure point (Legg, 2021) for communities to engage with formal actors but can also directly contribute to collective efficacy. Given the close alignment of the third sector with communities, they can utilise informal mechanisms of collective efficacy, similar to those employed by citizens, but with the added support of organisational resources.

Examining the mission statements and guidance practitioners provide is useful to gain insight into how this can be achieved. The "Charities Against Hate" (2020) campaign<sup>4</sup> is a notable example in the UK. Although their initiative was primarily developed to address online hate, the recommendations outlined in the guidance

---

<sup>4</sup> The "Charities Against Hate" campaign is a collective of over 40 UK-based charities working together to mitigate tensions.

apply to both online and offline contexts and to tensions more broadly. As such, many of their key objectives align with mechanisms for achieving collective efficacy (Sampson et al., 1997). Broadly, their guidelines can be divided into three main objectives: (i) engaging partners and widening the network to other agencies, (ii) challenging hate, and (iii) promoting positive, diverse, and inclusive content. The first objective is enhancing social capital within communities and among external partners through building multi-agent networks and providing support systems. NGOs understand the importance of “collective work and the creation... of new ways for people to come together and meet in different and relational spaces” (Kagan & Duggan, 2011, p. 401).

Conversely, the latter two objectives align with informal mechanisms for collective efficacy, such as promoting cohesion and trust and informal social controls. The recommendations of the campaign expand upon the objective of challenging tensions by stating that third sector groups should “visibly” address antagonistic behaviours in a “respectful, authentic, unapologetic” manner (Charities Against Hate, 2020). This objective aligns with informal social controls, which refer to intervening when something threatens a community (Sampson et al., 1997).

Furthermore, the guidance specifies that such challenges should be visible and made public, which also aligns with the concept of informal social controls, as it signals to others that the behaviour is unacceptable (Groff, 2015, p. 90). This approach establishes norms and values that reject divisive behaviour, ultimately promoting collective efficacy within communities. Condemning them, rather than responding to it with force, can be more effective in promoting collective efficacy, as it allows for changes in perceptions to occur organically rather than artificially suppressing them (Levin & McDevitt, 2002).

These processes closely align with the concept of “direct informal social control,” which refers to direct intervention by citizens, or in this case, NGOs. Alongside “direct informal social control,” Warner (2007, p. 99) also identified the concept of “indirect informal social control,” which refers to when citizens summon formal authorities, such as the police. The Charities Against Hate (2020, p. 8) guidance also reflects this by advocating for networking with public agencies and reporting hate,

when possible, which can be seen as a form of "indirect informal social control". The final recommendation of the Charities Against Hate report, "promoting positive, diverse, and inclusive content," aligns closely with the concept of "developing mutual trust" and promoting "cohesion", as proposed by Sampson et al. (1997).

This strategy of promoting positive and inclusive content is based on the idea that exposure to diverse cultures and perspectives can increase understanding and reduce tensions among different groups. This aligns with the traditional concept of contact theory, as Allport (1954) outlined, which posits that positive interactions with outgroup members can lead to decreased tensions and improved intergroup relations. This idea was later expanded on, suggesting that positive contact through knowledge-based campaigns can raise awareness and reduce negative perceptions of outgroups (Dovidio et al., 2017). Mergel (2013) asserts that a continuous supply of relevant new information should guide the operations of organisations. In this context, communities can be considered a sensor, where their active engagement and participation can enhance third sector operations.<sup>5</sup>

### ***2.3.6. Collective Efficacy in Digital Society***

The advent of digital society has profoundly impacted collective efficacy (Costello et al., 2017). In digital societies, the nature of community is transformed. Castells (1997) argues that communities are no longer organised based on spatial or temporal dimensions in the digital age. They now exist in digital spaces and differ from their offline counterparts in size and composition (Dutton, 1996). Scholars agree, however, that communities also need to restrict online transgressions (Oskuie et al., 2021), where transgressions entail "(un)conscious overstepping of moral or legal boundaries that confronts (un)written rules" (Hermes & Hill, 2021, p. 4).

---

<sup>5</sup> NGOs have been known to employ these strategies regularly, but little academic research has been conducted to examine the specific strategies and tactics used in these campaigns. This lack of research is addressed in Chapter Six, which aims to explore how positive and inclusive content is promoted, particularly in digital spaces.



Returning to Sampson et al.'s (1997) distinction between informal mechanisms and formally (externally) induced actions, it's crucial to emphasise that the latter often apply restrictions to online behaviours derived from offline contexts, which can result in errors, primarily because online environments possess their unique dynamics and challenges (Costello et al., 2017). In particular, formal stakeholders such as social media companies have faced criticism for their inadequacies in effectively restricting cyberhate and antagonistic speech, attributed to insufficient monitoring, inconsistent enforcement of policies, and prioritisation of free speech concerns over curbing online transgressions (Hawdon et al., 2015). Consequently, scholars have shifted their research focus towards informal mechanisms as a means of fostering collective efficacy on social media platforms, (Kim, 2015; Ireland et al., 2020; Glassman et al., 2021), recognising the limitations of external, formal controls in mitigating the complexities of tensions and its impact on digital communities.

Most studies on the impact of informal controls on collective efficacy have focused on measuring informal social controls (Costello et al., 2017; Hall, 2018; Hawdon et al., 2018; Ozalp et al., 2020). As previously stated, informal social control can manifest as direct or indirect interventions (Warner, 2007). This applies to digital contexts also, where indirect interventions occur when individuals report cases of cyberhate to external authorities (Ireland et al., 2020), such as the police or social media moderators, while direct interventions occur when individuals directly challenge antagonistic speech or cyberhate in digital spaces (Bastiaensens et al., 2014). Although most research has focused on examining direct interventions in digital spaces (Ozalp et al., 2020), the questionnaires and Likert-scale methods used to collect the data have failed to show a significant association between increased collective efficacy and reduced victimisation (Costello et al., 2017).

These limitations have led researchers to shift from traditional methods (e.g., interviews and surveys) to adopting social media data to comprehend information dissemination and flow (Williams & Burnap, 2016; Ozalp et al., 2020). Public endorsement on social media is evaluated by utilising platform-generated numerical metrics such as likes or retweets of a post (Boyd et al., 2010). These techniques were initially used in digital tensions contexts to investigate which public actors received the most endorsement after trigger events. However, they have recently

been employed to specifically examine how counter-speech contributes to upholding collective efficacy (Ozalp et al., 2020). Retweet rates were used to compare public endorsement between those spreading antagonistic sentiments, those offering counter-speech, and digital bystanders. The results provided the first empirical evidence that direct forms of informal social control can lead to increased collective efficacy in digital spaces (Ozalp et al., 2020).

This approach is limited by only assessing direct interventions, such as counter-speech, and cannot detect indirect interventions. Thus, the analysis in this thesis builds upon these recent approaches (Ozalp et al., 2020) by applying information diffusion analysis of engagement metrics to direct informal social controls in two ways. Firstly, NGO engagement rates are analysed when they participate in counter-speech, compared to their general day-to-day tweets. Secondly, the study aligns more closely with the work of Ozalp et al. (2020), where counter-speech is evaluated after a trigger event to determine how much collective efficacy influences local micro-events, which have not been explored in prior literature. Prior research has mainly focused on evaluating the presence or effectiveness of informal social controls (Sampson et al., 1997; Wells et al., 2006; Zhang et al., 2007; Twigg et al., 2010; Costello et al., 2017; Ozalp et al., 2020). However, there is limited research (with the exceptions of Pennant, 2005 Putnam, 2007 Livingston et al., 2008; Twigg et al., 2010) on how mutual trust and cohesion are strengthened in daily contexts. Understanding the formation of environments characterised by mutual trust and cohesion is critical, as they are significant factors in determining the involvement of individuals in informal social controls (Sampson et al., 1997; Simmons et al., 2005).

As previously noted, communities with high levels of cohesion are typically associated with traits such as "close-knitness," "mutual trust," and "shared values" (Sampson et al., 1997: 920). It is generally acknowledged that high cultural heterogeneity can make it challenging to achieve community cohesion. Established ingroup members may be less likely to participate in communal activities due to the perceived increase in outgroup populations and threats (Goodhart, 2004). Blalock (1956) first characterised these processes as inter-group conflict, highlighting the detrimental impact that increases in heterogeneity can have on community cohesion. Nevertheless, building on Allport's intergroup contact theory, scholars argue positive

exposure to different cultures can enhance feelings of trust, tolerance, and cohesion and contribute to a unified community with shared norms and values (Ferlander & Timms, 1999). NGOs are well positioned in this regard, as they are closely tied to community groups (Kelly, 2007) and have the opportunity to engage in informal mechanisms of collective efficacy using their resources. This thesis contributes to the literature by further exploring the role played by NGOs as knowledge campaigners (Perry et al., 2016) in promoting knowledge and awareness, thereby fostering cohesion and trust, and achieving collective efficacy within digital communities.

## **2.4. Digital Communication Strategies**

### ***2.4.1. Introduction to Communication Strategies***

As noted earlier, the rise of a worldwide digital society is marked by the marrying of digital technology and culture (Lindgren, 2017). Wildemeersch and Jütte (2017) suggest that social life has been reorganised around digital communication and media infrastructure. Such technologies have enabled individuals and organisations to transcend spatial and temporal confines when interacting (Castells, 1997). This shift has given unprecedented access to instantaneous high-quality content and opportunities for networking globally (Betro et al., 2012). This hasn't always been seen positively, with scholars noting new social threats in the digital society (Hawdon et al., 2015). As the digital society continues to reshape various aspects of social life, the subsequent discussion in this section concentrates specifically on the adoption of communications. The ultimate objective is to explore how these communication strategies can address emerging and pre-existing threats. Although most academics focus on how communication developments have altered social structures and individuals (Niedzwiecka & Pan, 2017), sociologists have also explored how they impact organisations.

Since individuals and social structures now operate and interact in digital spaces (Van Dijk, 2012), many organisations have seen the need to follow suit. Integrating social media into day-to-day operations can prove advantageous to organisations

because it helps, diversifies, and improves social connections with the general public and other stakeholders (Jue et al., 2010). As such, social media presents organisations with an opportunity to expand their audience reach, a process Castells (2009) called mass self-communication. This has enabled access to instant, high-quality content, and global networking opportunities (Betrot et al., 2012). Currently, most research on collective efficacy in digital society has centred around informal social controls and has focused on responses to specific incidents (Earl et al., 2013; Costello et al., 1997). As previously discussed, NGOs hold a unique position as a blend of public institutions and communities (Kelly, 2007: 1010). As such, they can advocate for policy changes while contributing to informal collective efficacy (Guo and Saxton, 2010). Accordingly, NGOs endeavour to foster community cohesion through several means such as (i) actively engaging partners and communities, (ii) counteracting hate by proactively challenging it, (iii) promoting positive, diverse, and inclusive content, and (iv) by continuously gathering information to help inform their efforts.

The initial objective prioritises enhancing social capital with community members, digital audiences, and external partners, including law enforcement agencies and governing entities. This is deemed crucial in digital realms, as these environments facilitate an increase in both the quantity and quality of connections. As a result, organisations in the third sector strive to augment collaborative efforts within these "relational spaces" to reinforce social capital among communities and stakeholders (Kagan & Duggan, 2011, p. 401). Studies have established that social media is a competent tool for organisations to engage current and potential stakeholders within multi-agency settings, strengthening collective actions and information-sharing (Greenberg & MacAulay, 2009; Guo and Saxton, 2010) and that the enhancement of social ties has a positive impact on advocacy efforts and is essential for informal mechanisms aimed at achieving collective efficacy by improving social capital within communities. The final two objectives pertain more specifically to the informal mechanisms of collective efficacy based on digital "netiquette" (Shea, 1994). Although interconnected, they differ in their specificities. The second objective emphasises the role of NGOs in engaging in informal social control practices, such as utilising counter-speech. The third objective is the central theme of Chapter Six's analysis. It aims to promote knowledge and awareness to foster cohesion and trust,

ultimately achieving collective efficacy. Scholars have identified social media and digital content as a useful tool for organisations to address issues associated with community tensions (Chakraborti, 2018: 400). As such, the day-to-day third sector promotion of inclusion and tolerance is the object of study rather than responses to specific tensions and events. A further understanding of gradual cohesion building is especially relevant in the digital era, given the expansive potential for information diffusion (Van De Velde et al., 2015) and NGOs' strong connections with community groups (Kelly, 2007).

This section situated this research within the broader framework of collective efficacy accomplished through formal and informal actors in different spaces. Building on this foundation, the current section focuses on NGOs' communication strategies and tactics to promote greater tolerance, inclusiveness, and positivity in digital spaces. This section provides a comprehensive overview of organisations' communication strategies and tactics, guided by Mergel's (2010) digital communication typology. More specifically, initial evidence suggests that the overall strategies employed in digital spaces can differ significantly depending on the key objectives set by organisations. Broadly, objectives reflect what sector (i.e., private, public and third) organisations operate in. However, they can be further stratified within these. This thesis considers two organisational uses of social media—first, the multi-agency use of instant messenger RTC platforms in the Welsh cohesion delivery network. Second is how NGOs employ public-space application communications to try to curtail tensions and uphold collective efficacy. This section focuses on the latter. Given the specificity of this topic and the limited surrounding literature, an understanding of digital communication strategies at a more general level is needed.

#### ***2.4.2. Digital Communications in the Third sector***

Scholars have paid less attention to how the third sector has utilised social media platforms than the public or private sector (Guo & Saxton, 2010). This could be attributed to NGOs having more recently adopted digital technologies than their counterparts (Betrot et al., 2012). Despite this, the turn of the millennium saw significant upticks in NGOS social media usage (Briones et al., 2011). This has aided

NGOs to bolster how they serve clients, engage donors, and communicate with staff and volunteers (Kaplan & Haenlein, 2009). Despite this, no single strategy has been found that is unanimously employed amongst NGOs, with many failing to incorporate its use in their operations (Deschamps & McNutt, 2014). Perhaps no cohesive communication strategy is expected to be evident amongst all NGOs, given that many have starkly different objectives (Curtis et al., 2010). This is akin to the wide array of public sector organisations discussed earlier in this chapter.

Initial studies have drawn on fundamental differences by categorising NGOs into three groups: education, environment, and disaster relief (Deschamps & McNutt, 2014). Although analysis of such categories has provided crucial insights into how NGOs have incorporated digital communications, they fail to recognise other organisation types. A notable category omitted is pro-minority rights organisations as a means and promote collective efficacy. This thesis, therefore, fills a gap in the literature by focusing specifically on these, with data analysis in Chapter Six exploring how pro-rights organisations in Wales use social media communications. Understanding this is important, given social media's apparent transformative potential to foster social inclusion and build community participation (Vance et al., 2009).

### **2.4.3. Strategy Frameworks**

Communication strategies broadly refer to content and presentation of social media activity and are often driven by a desired set of outcomes (Mergel, 2010). Although the benefits of using online communications are apparent, surrounding literature has shown that harnessing and implementing an effective strategy has proved challenging for private (Castronovo & Huang, 2012), public (Meijer et al., 2012) and third (Purdy, 2011) sector organisations. This has often been attributed to a need for more expertise in digital communications among social media administrators in such organisations (Water et al., 2009). Another reason for organisations failing to optimise social media usage is the lack of standardised smart practices and success criteria in the surrounding literature (Walting, 2011). Traditionally, organisational strategies are regarded as specified methods and tools to achieve predetermined

goals (Porter, 1980). However, this perspective has been challenged, with many arguing that strategies can develop when partially or entirely unplanned (Chaffee, 1985). Mintzberg (1978) advocated for strategy definitions to focus on patterns in streams of decisions instead. Strategies are only sometimes formally codified in documents but can instead develop due to organisation members' choices (Mergel, 2012). Unplanned strategies were later coined as emerging strategies, with planned ones labelled deliberate ones (Mintzberg & Waters, 1985).

Given that no standardised approaches have yet been made in the literature consensus (Walting, 2011), a closer consideration of emerging strategies is pivotal. Moreover, many digital communication strategies emerge organically through decisions made by individuals working in the organisations (Mergel, 2012, p. 290). Despite this, organic strategies can still be categorised and identified, regardless of the degree to which they were consciously designed. Mergel (2010) proposed four broad digital communications strategies: push, pull, networking and transactional. Mergel (2013:327) explains that organisations can 'trace their online interactions' to define their 'mission support' and overall social media tactics using this framework. In other words, it helps to categorise different content types published online. Mergel's (2010) typology offers an appropriate framework for analysis for this thesis because it provides a standardised base point for analysis that can be applied to minority rights NGOs. Given that Mergel's (2010) strategies are so broad, it is no surprise that they differ significantly.

#### *2.4.3.1. Push Strategy*

Push strategies utilise social media platforms as broadcasting channels, where organisations communicate with audiences using one-way content (Mergel, 2012). These approaches are more closely aligned with those exclusive to Web 1.0 and operate in a broadcast paradigm where organisations do not seek to interact directly with audiences (Greenberg & MacAulay, 2009). In this context, push strategy approaches post content on social media pages to enhance content diffusion by bolstering information reach (Van De Velde et al., 2015). Push strategy communications reflect those made by the media, such as television, newspapers,

and radio outlets (Mergel, 2012). Push strategy communications have proved advantageous for organisations because they can similarly broadcast information to media sources without relying on them. As such, organisations provide information to the public that may not be deemed 'newsworthy' but is still essential. This is arguably more relevant from a third sector perspective than the public sector (Mawby, 1998).

Conversely, the day-to-day operations of NGOs, mainly those not involved in disaster relief, are less frequently covered by the media. Push strategy communications provide NGOs with unprecedented opportunities to broadcast unique content whenever they want in digital society's round-the-clock media culture (Goldsmith, 2015). In doing so, organisations can build an online presence with greater reach using information dissemination (Van De Velde et al., 2015). This helps them communicate with pre-existing stakeholders while also helping them access broader audiences, leading to greater exposure and awareness (Mergel 2013). Digital communications, therefore, enable organisations to push informational and educational content to a much wider audience (Van De Velde et al., 2015).

Push strategy communications are also employed to promote transparency, accountability, and trust (Mergel, 2010). This is a major priority in public sector organisations, particularly those under frequent scrutiny. For example, police digital communications research found that a prevalent motivation for using push strategy tweets is 'image building' (Mawby, 2010; O'Connor, 2015; Kappeler & Gaines, 2020). In the public sector, most image building is accomplished by outlining organisational successes (Kappeler & Gaines, 2020) and providing transparency about day-to-day operations (Mawby, 2010). However, image-building activities have proved challenging in the public sector because they open organisations to greater scrutiny than before (Goldsmith, 2010). Organisations must address frequent criticisms in the public sphere from anonymous sources (Schneider, 2014). Although some merit has been found in image-building techniques, many public organisations have suffered from the overly professional tones they use while communicating. Professional tones can make organisations seem less relatable in digital environments and disfranchise audiences (Brainard & McNutt, 2010).



On the one hand, they want to promote engagement with their audiences to promote transparency and build good public images. However, on the other, they need to maintain a degree of formality in their work, particularly when addressing serious matters such as murders, fires, and kidnappings. Although formal tones are evident in public sector push strategy communications (Brainard & McNutt, 2010), little is known about what tones are used for NGOs. Despite this, some studies have shown that NGOs likewise use push strategies to promote transparency (Deschamps & McNutt, 2014). However, less evidence exists to show these are designed to curtail negative feelings towards them but instead to include and update other external stakeholders (Deschamps & McNutt, 2014).

Initially, Web 1.0 functions only enabled organisations to use one-way push strategies within a broadcast paradigm (Greenberg & MacAulay, 2009). The emergence of Web 2.0, and more specifically, social media, has given way to further methods of communication. Perhaps most notable is the greater opportunity for audience and stakeholder engagement. This has been characterised as the emergence of the dialogical paradigm (Greenberg & MacAulay, 2009). It is important to note that the dialogical paradigm has not replaced the broadcasting paradigm but exists alongside it. Many organisations still employ one-way push strategies for reasons detailed earlier in this section. In contrast, others employ two-way communication strategies to develop strong ties. The degree to which organisations want to waken or bolster their ties to stakeholders and audiences often reflects their key objectives and missions (Mergel, 2013). Despite this, scholars note that all organisations seek, to some extent, to widen and strengthen social capital in their networks (Kagan & Duggan, 2011)

Different NGO types have been found to employ a range of communication strategies. Disaster relief organisations utilise dialogical strategies that offer two-way communications to build intense or strong relationships with audiences and associated stakeholders (Deschamps & McNutt, 2014). Conversely, educational organisations employ one-way broadcasting techniques associated with Mergel's (2010) push strategy. These provide an informational-driven means to build weak ties based on common interests (Deschamps & McNutt, 2014). Environmental organisations were found to employ a mixed approach, with many examples of

information broadcasting alongside dialogical two-way communications. Differences between NGO approaches are expected, particularly when considering their starkly different objectives.

Interestingly, organisations with similar purposes may employ vastly different communication strategies. For example, a study that explored police communication strategies amongst Canadian police forces found that an array of conflicting strategies were used from force to force (O'Connor, 2015). This included push-dominated broadcasting approaches, dialogical-driven pull and networking strategies and merged approaches using Mergel's key strategies (O'Connor, 2015). As mentioned, surrounding literature has failed to include minority rights NGOs in such comparisons. The approaches made by such organisations are currently, therefore, unknown. Analysis in Chapter Six sheds light on how minority rights organisations compare strategies to their disaster relief, educational and environmental counterparts. Dialogical communications that bolster audience and stakeholder ties (Granovetter, 1973) relate to two of Mergel's (2010) communication strategies- pull and networking. Indeed, dialogical approaches bolster pre-existing ties with stakeholders and help expand and diversify the organisation's networks to include a wider range of internal and external inputs (Jue et al., 2010; Kavanaugh et al., 2012; Ma, 2013).

#### *2.4.3.2. Pull Strategy*

Although pull and networking communication strategies are dialogical, they differ in how they actively engage audiences and stakeholders. Pull strategy communications employ a two-way approach in which organisation audiences are actively encouraged to participate (Mergel, 2013). These participations, however, are not encouraged directly on social media platforms but rather to other forms of engagement. Examples of pull strategy techniques include participation in surveys, focus groups, links to organisation-run web pages, appeals for information and recruiting new staff or volunteers (Mergel, 2012; O'Connor, 2015).

Moreover, pull strategy approaches are more closely aligned with the dialogical paradigm than the broadcasting paradigm (Greenberg & MacAulay, 2009). However, Mergel (2012) notes that organisations rarely respond to comments made on these posts. In other words, they seek to pull audiences to other forms of engagement, not on social media. These still lead to strong social ties between audiences and organisations but do not occur exclusively in digital spaces (Granovetter, 1973). Pull strategy tweets promote engagement and participation in which organisations use followers and stakeholders as sensors for information or general input (Mergel, 2012). By increasing information depth, pull strategies can promote content diffusion (Van De Velde et al., 2015). Some platforms, such as Twitter, have restrictions on text length (Han et al., 2019), which can limit the depth of content. To overcome this, organisations may use pull strategies to allow their audience to engage with more extensive content published by the organisation on external websites (Mergel, 2010). Conversely, initial evidence from the third sector indicates that pull strategies are predominately used to elicit volunteers, engage civil society, and consult public opinions on issues using surveys and focus groups (Deschamps & McNutt, 2014). Dialogical pull strategies, therefore, are advantageous in integrating a wider spectrum of internal and external stakeholders into the day-to-day operations of NGOs (Jue et al., 2010; Kavanaugh et al., 2012; Ma, 2013).

#### *2.4.3.3. Networking Strategy*

Networking communication strategies are similar to those detailed in the two-way pull approaches. They also seek to strengthen social ties with audiences and key stakeholders by providing wider opportunities for engagement and expanding networks to a wider range of internal and external stakeholders (Granovetter, 1973; Jue et al., 2010; Bertot et al., 2010; Hrdinová et al., 2010; Kavanaugh et al., 2012; Ma, 2013). However, networking interactions are more specifically confined to social media platforms (Meijer & Thaens, 2013). When theorising communication strategies, Mergel (2010, p. 10) realised social media's potential for bridging organisations and public audiences: 'the use of social media tools is highly interactive with much back and forth between the agency and its diverse constituencies'. Instead of implementing a two-way pull approach, in which public

audiences and external stakeholders are used as information sensors, the networking strategy promotes a two-way cross-boundary interactive framework (Mergel, 2013). Organisations network with public audiences, resulting in an interactive co-design of services in which community building and the establishment of 'issue networks' are made possible (Mergel, 2013). Networking communications are, therefore, more direct, and interactive than pull and push strategies, as such organisations employ networking strategies to engage in direct dialogue with audiences on social media platforms.

This helps bolster social capital with community members and external partners by building relationships (Bortree and Seltzer, 2009; Lovejoy and Saxton, 2012). Although pull communications are still located in the dialogical paradigm, networking strategies surpass these and can be realised as the purest form of dialogical interaction (Greenberg & MacAulay, 2009). Broadly, networking communications are achieved by asking questions or polls, replying or commenting to external posts and directly messaging individuals (Mergel, 2013). The latter interaction has, however, rarely been included in studies that employ Mergel's framework because of privacy regulations associated with collecting social media messaging data (Meijer & Thaens, 2013; O'Connor, 2015). Considering this, it is reasonable to assume that many cases of networking communications are not reported in strategy prevalence data. This presents real issues to researchers because some of the most meaningful interactions between organisations and their constituents are likely to occur in private messages.

Another way in which networking strategies depart from push and pull is in their tasks and key objectives. Push and pull communications are strategies with centralised tasks (Meijer & Thaens, 2013). Both approaches have broad objectives that are consistent throughout most cases. In push, these relate to providing information and education while also promoting transparency to help build a positive image (Mergel, 2013). For pull strategy communications, central tasks include using the public as information, resources, and sensors through non-social media consultations and interactions (Mergel, 2013). In both cases, the broad objectives are easily identifiable and have been found to manifest in easily distinguishable sub-categories (O'Connor, 2015). Networking communications, however, have been

identified as decentralised (Meijer & Thaens, 2013). This means that the services and tasks completed in networking communications relate to a much wider variety of objectives that are often unique each time and reflect the co-design of services promoted in the approach, in which organisation constituents have an equal say in the interactions (Mergel, 2013).

#### *2.4.3.4. Transactional Strategy*

The final communication strategy outlined by Mergel (2010) is transactional. This involves organisations utilising online communications to transact with external stakeholders or audiences regarding physical goods or services (Mergel, 2012). Transactional strategies are prevalent among private sector organisations using online communications as a marketing tool to sell products and services (Thomas, 2007). In contrast, transactional approaches are less common in the public sector, where push, pull, and networking strategies are more widespread (Mergel, 2012; Meijer & Thaens, 2013; O'Connor, 2015). In the public sector, transactions reflect public services provided to communities or constituents, with organisations using transactional strategies to enhance these services (Mergel, 2012). Scholars have consistently emphasised the significance of organisations providing support services to marginalised groups.

Nevertheless, concerns persist about the limited accessibility of these services, indicating that only a privileged few, often those in more economically or socially stable circumstances, are aware of such resources. Notably, even among these individuals, a significant portion remains uninformed about these services (Hardy and Chakraborti, 2020:144). There is a paucity of literature on social media's impact on providing support services to constituents. Therefore, this presents an opportunity for analysis in Chapter Six to examine the potential impact of social media on providing support services to marginalised groups. Social media is described by Mergel (2012) as a transactional environment where constituents are viewed as business partners of public agencies.

NGOs commonly use social media to solicit donations (Deschamps & McNutt, 2014). Donations were traditionally made through general appeals, but non-profits also use transactional incentives such as prize giveaways, trivia contests, and product sales (Castillo et al., 2014). NGOs use transactional communications more frequently than the public sector (Deschamps & McNutt, 2014), likely due to the longstanding tradition of donations to non-profits and the emergence of transactional approaches in public sector services (Mergel, 2012). However, third sector transactional approaches are less common than market-driven communications in the private sector (Thomas, 2007). Donation appeals also vary between NGOs, with some, such as disaster relief non-profits, appealing at a higher rate than others, like educational organisations, that broadcast informational content on policy (Deschamps & McNutt, 2014). The extent of transactional approaches used by minority rights non-profits is unclear. Chapter Six provides insights into the prevalence of transactional approaches among minority rights organisations compared to other strategies.

#### ***2.4.4. Content Tools and Modalities***

Other than communication strategies, organisations use content tools or modalities when posting on social media platforms. Content tools do not only refer to literal text, such as language tones or word counts, but also URLs, hashtags, mentions and emojis (Han et al., 2019). Much like communication strategies, patterns of use for content factors can develop partially or entirely unplanned. These reflect individual decisions made by members of organisations that organically develop rather than being strategically pre-determined (Chaffee, 1985). A major drawback of organically implementing content tools is that organisations have no standardised smart practice or success criteria (Walting, 2011). Despite this, a standardised model is perhaps unattainable, given the vastly different objectives between organisations and organisational sectors. Mentions on social media platforms can be broadly categorised into two forms: user mentions and replies (Ahmed, 2018). User mentions refer to when a tweet mentions another user without any prompt using the @ symbol.

In contrast, replies are when a user is mentioned using the @ symbol in response to a tweet in a thread. Despite the distinction between these two mentioned forms, it can be challenging to differentiate them in raw data, leading to most studies merging the two categories (Anger & Kittl, 2011). Mentions and replies are primarily used to instigate open dialogues with external organisations or the general public (O'Connor, 2015). This dialogue-oriented approach aims to build cybercommunities by strengthening ties and expanding networks. As such, mentions are most closely aligned with networking strategies (Mergel, 2010). Non-reciprocal friendship structure platforms such as Twitter facilitate this process by allowing users to mention others they do not follow (Davenport et al., 2014). Interested users can access these mentions and subsequent replies by viewing threads (Mergel, 2012). As such, mentions serve as networking tools that prioritise strengthening existing ties and creating new ones over disseminating information (O'Connor, 2015).

As previously noted, other content tools, such as hyperlinks and hashtags, focus more on information propagation and diffusion than strengthening ties (Van De Velde et al., 2015). Although both tools enhance information dissemination, how they achieve this differs significantly. Hyperlinks serve as a tool for providing greater detail on information to audiences by linking to content from other organisations (Van De Velde et al., 2015). This can include "longer blog updates or links to other media, such as press releases on official government websites" (Mossberger et al., 2013, p. 355), which allow for more in-depth discussions on topics and bypasses the textual limitations of platforms like Twitter (Han et al., 2019). These links often represent pull strategies that aim to further engage social media audiences with organisation-published content (Mergel, 2010), thus providing more information than is possible on Twitter and integrating audiences more deeply into the organisation's work (Berman et al., 2007; Walters et al., 2009). This can also contribute to developing ties with audiences, particularly when the content provides transparent accounts of the organisation's operations (Park et al., 2002; Segerberg & Bennett, 2011). Hyperlinks can also redirect audiences to external sources of information, which, to some extent, aid in developing networks using the public endorsement of other organisations.

Using hashtags on social media platforms strengthens information diffusion by expanding the reach of information. Hashtags are a form of metadata tags and are characterised by the use of the "#" symbol. According to Mergel (2012), hashtags are widely used on social media platforms to enable users to search for specific topics. The primary function of hashtags is to increase the reach of tweets by increasing their searchability, as Zappavigna (2011) noted. This, in turn, can significantly increase audience reach, exposing many individuals to new content, especially if it is related to a trending topic (Mergel, 2012).

Yet, despite their uniform appearance, hashtags can serve various purposes depending on their intended audience and usage. The utilisation of hashtags has been extensively studied in both the public and private sectors, with research highlighting how they are employed to achieve various evolving objectives (Kleinberg, 2013; Ayanso & Moyers, 2015; De Widt & Panagiotopoulos, 2018; Han et al., 2019). In private contexts, hashtags are commonly utilised by companies as a means of "supplementing their brand images" by fostering a sense of community among customers, which in turn leads to increased engagement (Kleinberg, 2013: 3). On the other hand, public actors have been observed to use hashtags as a means of disseminating government information more widely and building relationships with the public (De Widt and Panagiotopoulos, 2018: 2). Despite the distinct goals of these sectors, both utilise hashtags as a means of building social capital with their audience and expanding their reach. Using hashtags to broaden information channels is an example of "mass self-communication" (Castells, 2009).

NGOs' use of hashtags on social media has received limited scholarly attention. According to Bruns and Burgess (2011), hashtags have the potential to effectively nurture long-term communities of tolerance through the promotion of knowledge, awareness, and social capital. Therefore, understanding the usage of hashtags by NGOs is crucial to community cohesion. Saxton et al. (2015) categorised hashtags used by NGOs. This typology, consisting of eight self-explanatory forms, is a valuable framework for research in this field. These include "public education," "events," "call-to-action," "values and goals," "branding," "dialogic," "time and place," and finally, "business hashtags". (Saxton et al., 2015)



Aside from hashtags, emojis have been utilised as content tools on social media platforms. They are "small icons depicting objects or emotions often used in personal instant messaging and social media contexts" (Zhang et al., 2022, p. 618). Despite the growing popularity of emojis, their usage in organisational contexts on social media is not yet common (Waters and Jamal, 2011; Casado-Molina et al., 2019; Zhang et al., 2022), with academic consensus suggesting that NGOs may be more likely to implement their use due to their close alignment with citizen use of social media (Chen, 2019). Despite this recognition, there is limited knowledge about how NGOs use emojis. In light of this, Chapter Six provides an exploratory analysis that fills this gap by examining the types and contexts of emojis utilised.

Current evidence suggests that emojis are often utilised in dialogical forms, such as in replies rather than original tweets or retweets (Kwon & Sung, 2011). Emojis have also been found to increase an organisation's relatable nature among community groups (Hayes et al., 2020), aligning with dialogical strategies to strengthen and broaden organisation networks through positive image building (Mergel, 2010). In this sense, emojis have helped overcome the lack of non-verbal cues on social media, allowing organisations to communicate specific emotions to audiences (Bai et al., 2019). The use of emojis to promote diversity, inclusivity, and solidarity has been explored in a growing body of literature. These efforts have taken the form of informal social controls in response to specific events, such as the BLM protests following George Floyd's death (Santhanam et al., 2021), and promoting inclusivity on a day-to-day basis. The Diversity Language Model (Swartz et al., 2020, p. 95) offers a useful typology in this regard, positing that emojis can represent minority groups and can be categorised into four main categories: race, religion, gender, and sexual orientation (Swartz et al., 2020, p. 95).

In the previous section, it was highlighted that the digital society has brought about a transformation in the way collective efficacy is measured, moving away from the use of survey data, such as Likert-scales (Sampson et al., 1997), and toward the use of social media engagement data (Ozalp et al., 2020). This approach offers a more precise method of evaluating the endorsement of specific sentiments compared to others. It is based on public endorsements measured by platform-generated engagement metrics, such as likes, retweets, and replies on Twitter (Boyd et al.,

2010). Using these metrics instead of Likert-scales is valuable in obtaining real-time or near-real-time information (Edwards et al., 2013). However, prior research has limited its application to instances of informal social controls, where users intervene for the common good (Ozalp et al., 2020). In this light, the analysis in Chapter Six extends the existing literature by investigating the use of engagement metrics in day-to-day contexts of inclusivity promotion. The aim is to use engagement metrics on Twitter to compare various communication strategies and content tools to determine which are most effective in supporting central organisational goals, such as engaging partners, broadening networks with other agencies, counter speech, and promoting positive, diverse, and inclusive content.

## **2.5. Trigger Events for Community Tensions**

Community tension research has traditionally focused on the spatial distribution of incidents, examining factors such as spatial clustering, demographic characteristics, and economic causes (Braga et al., 2012). While zonal analyses can provide valuable insights into underlying patterns, many scholars argue that the temporal dimension of these tensions has been disproportionately overlooked in the literature, with an increasing number of studies advocating for further examination of the temporal dynamics (King & Sutton, 2013; Legewie, 2013; Hanes & Machin, 2014; Williams and Burnap, 2016; Devine, 2018; Ozalp et al., 2020). Increases in academic attention to temporal analyses can be attributed to changes in available data. Namely, the emergence of big data was brought about by digitalising social life (Shahidullah, 2018). Work that focused more on spatial explanations of utilised data sources such as interviews, police-recorded data and surveys. While these punctiform methodological tools prove effective in outlining tensions and locations, they could be more effective in pinpointing when and how they develop.

When examining community tensions through a temporal rather than a spatial paradigm, scholars first used offline data sources to pinpoint when spikes occurred. Initial offline data sources included police-recorded statistics (King & Sutton, 2013; Hanes & Machin, 2014), interviews and survey data (Legewie, 2013). Using such data sources enabled researchers to identify temporal patterns of community

tensions for the first time. Subsequent temporal inspections revealed frequent yet non-periodic spikes (for example, see Williams, 2021, p. 125).

Since each spike occurred non-periodically, initial interpretations could allude to them being random. However, a reoccurring theme was present in every example- they were all preceded by an identified 'trigger event' Williams (2021: 215). While the term trigger event is widely used in related work (King & Sutton, 2013; Legewie, 2013; Hanes & Machin, 2014; Williams & Burnap, 2016; Devine, 2018; Ozalp et al., 2020), no single definition is provided in a community tension context. This thesis interprets trigger events as occurrences that galvanise and motivate negative attitudes or actions towards a minority group. This definition is purposely left broad because many studies only examine events exclusively relating to recorded hate crimes (e.g., King and Sutton, 2013).

### ***2.5.1. Underlying theories***

The understanding of the temporal clustering of community tensions in the wake of trigger events can be illuminated by examining various sociological and criminological theories. Such theories include strain theory (Cohen, 1955; Merton, 1957), social identity theory (Tajfel & Turner, 1979), intergroup theory (Stephan & Stephan, 2002), the justification-suppression model of prejudice (Crandal & Eshleman, 2003), stages of deviant events (Cohen, 1972), and intergroup crime and social control (Piatkowska & Lantz, 2021).

The relationship between intergroups was first explored by Blalock (1977), who posited that prejudice, and subsequent community tensions stem from a sense of threat perceived by the majority group towards a minority group in society. These notions of threat are an extension of what Cohen (1955) and Merton (1957) referred to as strain, which refers to the process in which crimes are committed as a response to a perception of instability or threat that is attributed to a designated group (Burch, 2022). In this context, prejudice-driven tensions are seen as reactive and driven by fears of socioeconomic security and encroachments upon the dominant group's identity (Hall, 2015).

Tajfel and Turner's (1979) social identity theory categorises individuals into two primary groups: ingroups (typically the "majority" group) and outgroups (often considered the "minority" or "designated other" group). This categorisation can lead to biases, stereotypes, and in-group favouritism, shaping attitudes and behaviours. Stephan et al.'s (2008) intergroup theory builds on this by identifying two ways in which the ingroup may perceive a threat from the outgroup: tangible threat, in which the ingroup feels threatened by potential harm or resource depletion from the outgroup, and symbolic threat, in which the outgroup's presence challenges the ingroup's self-image. This concept of tangible threat builds on Cohen's (2002) theory of predictive symbols of violence, which proposes that certain demographic traits may be perceived as indicative of the potential for violence. As a result, outgroup members are perceived as more threatening in such events due to the perceived risk of violence to the ingroup. Cohen (2002, p. xxiii) notes that such feelings can be bolstered by the traditional media, including newspapers, magazines (including tabloids and broadsheets), radio broadcasts, and television programming as forms of mass communication, 'setting the agenda' and 'transmitting the images. Tangible resources do not only include concepts such as land but also welfare spending made by Governments and access to public services (e.g., length of hospital waiting times attributed to migrants during the 2016 Brexit referendum). The second factor relates to a perceived threat to culture and views.

In this context, ingroups regard outgroups as a threat, drain or deterrent to their culture and way of life (Stephan & Stephan, 2002). It is important to note that threats and strains can be real or simply just perceived (Walters, 2011). As such, many threats and strains can be predicated on false information driven by derogatory ideologies (Sternberg & Sternberg, 2008). Threat and strain can, therefore, be identified as the first dimension of Crandal and Eshleman's (2003) JSM. JSM is hinged on the idea that perpetrators of bigoted attitudes or activities do so as a result of an ongoing psycho-social conflict (Williams et al., 2022) between "a desire to express an emotion and at the same time, maintain values and self-concepts that conflict with prejudice" (Crandal & Eshleman, 2003, p. 414). In this case, feelings of perceived threat and strain attributed to the outgroup serve as initial justifications for ingroup members to exhibit prejudiced behaviours.

Tensions arising from tangible and culturally perceived threats to the outgroup can be enhanced as the population size increases. As the ingroup's relative majority reduces, they feel an increased risk in their ability to sustain economic, political, and cultural dominance in society (Piatkowska & Lantz, 2021). As a result of this increased threat, Blalock (1977) suggests that the ingroup elicits prejudice. While Blalock and Stephan provide foundational insights into how inter-group tensions develop in the first place, they say less about how they form into temporal clusters. Scholars have since proposed that tensions, follow suit when a perceived threat is drastically enhanced by galvanising events (Williams, 2021). This is hinged on the outgroup's perceived threat to social order (Piatkowska & Lantz, 2021). The events that lead to enhanced perceived threat levels are also known as “trigger events”. As trigger events can galvanise these perceived threats, prejudiced behaviours can be justified more easily in the aftermath of these occasions. Scholars have observed that incidents fuelled by bias not only serve as precursors but also function as indicators of broader underlying community tensions (Schweppe and Perry, 2022). Trigger events can intensify these underlying tensions between ingroups and outgroups, such as social threats (e.g., terror attacks) or economic strain (Hopkins, 2010; Legewie, 2013). Moreover, temporal surges in tensions following trigger events stem from a sudden escalation in perceived ingroup threat from the outgroup, leading to the justification of prejudice (Crandal & Eshleman, 2003).

Black's (1983) theory of crime as social control is also helpful in understanding how trigger events cause temporal clustering of prejudice and community tensions. Black (1983) argued that sometimes prejudice reflects a self-help standpoint in which perpetrators (usually from the ingroup) rationalise committing a crime in pursuit of a perceived justice from their perceived grievance. These criminal pursuits for justice can often be bolstered when a feeling of no legal alternative is present. Pursuing what is perceived as justice plays a significant role in shaping JSM, tilting it more towards the endorsement of prejudice rather than its mitigation (Crandal & Eshleman, 2003).

When exploring trigger events, it is helpful to consider notions of moral outrage and moral cleansing (Tetlock et al., 2000; Skitka et al., 2001). Following trigger events, these dynamics materialise as ingroup members unfairly target innocent individuals

from the outgroup solely based on their affiliation with the same demographic as those connected to the initial trigger event. Black (1983) aptly termed this phenomenon "collective liability." Had ingroup members exclusively sought justice from those directly implicated in the initial trigger, it might have prevented the emergence of widespread community tensions. However, collective liability engenders unwarranted and pervasive tensions between two groups not directly involved in the triggering incident. This concept aligns closely with moral outrage (Tetlock et al., 2000), wherein individuals experience intense negative emotions, such as anger and resentment, in response to perceived injustices or threats from outgroups. This emotional reaction can often lead to the rationalisation and even escalation of biases and hostile attitudes towards the targeted outgroup, a phenomenon known as prejudice justification.

The notion of collective liability finds support in the research of Lickel et al. (2006), who observed that members of victimised groups (linked to the trigger event) would engage in acts of aggression and violence against innocent outgroup members, driven by a sense of collective blame on behalf of their fellow in-group members (Piatkowska & Lantz, 2021). Lickel et al. (2006) coined this process "vicarious retribution." The theories of collective liability by Black (1983) and vicarious retribution by Lickel et al. (2006) have subsequently been embraced by scholars studying trigger events. Consequently, these two factors operate in concert to reinforce justifications for prejudice, particularly in the immediate aftermath of trigger events (Crandal & Eshleman, 2003).

Conversely, an opposing psychological process, moral cleansing (Skitka et al., 2001), can counterbalance this phenomenon. Moral cleansing is the process by which individuals, after witnessing or being involved in events that trigger bias or hatred, strive to alleviate feelings of guilt or moral impurity. They achieve this by taking actions consistent with their moral values or distancing themselves from the negative events or behaviours associated with bias and hatred. These actions can effectively counteract prejudice and foster an atmosphere of tolerance and inclusion. In situations where individuals perceive injustices or threats, moral outrage can ignite intense negative emotions and lead to the justification of prejudice. However, it's important to note that some individuals may also engage in moral cleansing to

assuage their guilt and restore their moral integrity. This can lead to actions that actively counter bias and promote tolerance among diverse groups. The intricate interplay between these psychological processes plays a pivotal role in shaping responses to trigger events and significantly influences their impact on intergroup relations.

### ***2.5.2. Trigger event stages.***

When considering the list of sudden peaks in tensions and possible subsequent hate crimes and incidents at the beginning of this section, Williams (2021, p. 215) assessed the existing literature to thematically identify four trigger events categories: (1) political votes, (2) policy changes, (3) court cases and (4) terror attacks. Although the categories are seemingly very different, short and intense spikes of community tensions have been identified in their immediate aftermath. Hanes and Machin (2014) outline this phase as a shock period whereby the public's attitude towards the suspected perpetrators (in the case of terror attacks) or associated groups or communities are significantly affected. Despite seeming most salient, the shock period following trigger events is not the only stage in their overall process. Understanding other stages before and after the shock period is important to understand how temporal clusterings of community tensions eventually subside (Ozalp et al., 2020). By considering Cohen's (1972) moral panics, we can identify four stages that are innate to deviant events (or series of events): (1) the warning stage, (2) the impact stage, (3) the inventory stage and (4) the reaction stage.

The warning stage occurs before the trigger event. Usually, pre-existing levels of perceived threat exist amongst the outgroup towards the ingroup. In the case of Woolwich, attack factors such as high terror threat levels coupled with previous attacks contributed to the heightened levels of perceived outgroup threat (Williams & Burnap, 2016). At this stage, justifications for prejudiced expressions exist but mostly do not outweigh suppressions (Crandal & Eshleman, 2003). The impact stage consists of the actual trigger (e.g., the stabbing of Lee Rigby in Woolwich). Social responses were highly disorganised during this period, with many social actors yet to internalise the trigger.

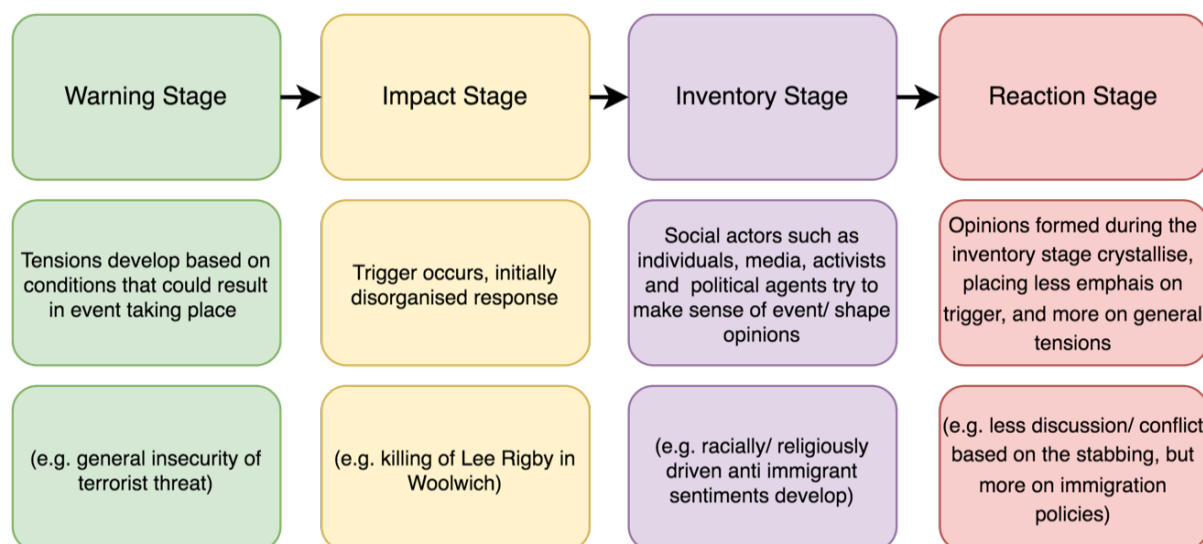


Figure 3: Cohen's (1972) stages of deviant events based on Williams and Burnap's (2016) trigger event case study (Woolwich terror attack)

Following the trigger comes the inventory stage, whereby social actors try to make sense of the event. Williams and Burnap (2016) note that during this stage, various social actors (i.e., the media and politicians) can play a role in shaping opinions by narrative building. Glaeser (2005, p. 3) described these influential social actors as 'entrepreneurs of hate' and argued that they usually benefit politically or economically from propagating inflammatory perspectives. Many deem these agents' impact as justification factors (or catalysts) for prejudice (Crandal & Eshleman, 2003). As such, a body of evidence and theoretical explanations have emerged, i.e., dangerous rhetoric theory, political legitimisation, and agenda-setting (Perry, 2001; Cohen, 2002; Koopmans & Olzak, 2004; Edwards & Rushin, 2018). The inventory stage is usually also when outgroup members become demonised, with additional notions of strain and threat being attributed to them. As stated earlier, these can be bolstered by pursuits of justice and collective liability (Black, 1983; Lickel et al., 2006).

Moreover, these stages apply to extremist triggers, e.g., terror attacks and the remaining trigger event types. Despite this, differences remain between the three commonly discussed trigger event types (political, extremist and court cases) and should be considered separately. The following section examines various trigger event case studies.



### **2.5.3. Extremist Trigger Events**

The most commonly explored trigger relates to extremist events (i.e., terror attacks) (Byers & Jones, 2007; King & Sutton: 2013; Hanes & Machin, 2015; Williams & Burnap, 2016; Müller & Schwarz, 2021; Czymara et al., 2022). Lending further evidence to Black's (1983) theory of collective liability and Lickel et al.'s (2006) theory of vicarious retribution, initial evidence suggested that terror attacks can result in increased levels of animosity towards members of the same racial or religious groups as perpetrators, resulting in increased tensions (Treadwell & Garland, 2011). As such, they become 'predictive symbols' of violence (Cohen 2002: xxiii). Since the 1990s, there has been an increase in global terrorist attacks, including the 9/11 attack on the World Trade Centre in 2001 and the 7/7 attacks on London's bus transport system in 2005. These events prompted academic inquiries into the relationship between terrorism and community tensions (King & Sutton, 2013). These studies found that after the 9/11 attack, 481 anti-Muslim hate crimes were recorded, with over half occurring at least two weeks after the event, indicating that temporal shocks in hate crimes can occur after large-scale terrorist attacks. Alarming, evidence from the Global Terrorism Index (2020) suggests that the rate of terror incidents is accelerating and creating clusters (see Sadique et al., 2018). Importantly, this clustered period allowed scholars to explore the relationship between terrorism and hate crimes (Piatkowska & Lantz, 2021).

Time-series data of such events provided initial evidence of temporal clusterings in community tensions. Later, regression models were devised to help explain whether these sudden increases can entirely be associated with the triggers (Piatkowska & Lantz, 2021) and which are built on King and Sutton's (2013) work by controlling for other factors, such as time lags, which denote the effect of Brexit. The first model revealed significant positive associations between the triggers and racially motivated hate crimes for all three events (London Bridge, Finsbury Park and Manchester Arena). The regression coefficients were strongest for the former two triggers (about racially motivated hate crimes). This reflected the initial findings made in the time-series analysis. The second model examined religiously motivated hate crimes and reflected findings made in the first. Once again, all three terrorist events were positively associated with religiously motivated hate crimes. One criticism of using

police recorded data is that it can fail to capture all aspects of hate in the wake of a trigger due to factors such as nonreporting and temporal granularity (Williams & Burnap, 2015; Schweppe et al., 2020). As such, this thesis builds on past studies by exploring community tensions more broadly moving beyond specifically just hate crimes and incidents.

#### ***2.5.4. Political Trigger Events***

The literature suggests that political factors can serve as catalysts for escalating community tensions. This has been observed to occur through two mechanisms: (1) political elections/votes and (2) policy changes (Williams, 2021). In light of the growing polarisation of Western politics, scholars have increasingly focused on the impact of political elections or votes on community tensions. King and Sutton (2013) posit that the mere occurrence of political elections or votes can lead to spikes in community tensions. Building on Blalock's (1956) theory of intergroup threat, Olzak (1990) found evidence that sudden changes in the political landscape can result in increased intergroup conflict due to a heightened sense of threat. This can occur through the perceived threat to the ingroup from the outgroup following shifts in political or cultural power (Piatkowska and Lantz, 2021) or through the legitimisation of discriminatory attitudes towards the outgroup by mainstream politicians, movements, or political parties (Perry, 2001; Koopmans & Olzak, 2004; Edwards & Rushin, 2018; Piatkowska & Lantz, 2021).

In the 2008 USA presidential campaign, the phenomenon of ingroup feeling threatened by a perceived increase in outgroup political power was demonstrated. As Barack Obama emerged as President, several visible tensions and subsequent incidents emerged (Williams, 2021). One notable case was the murder of Marcelo Lucero in Suffolk County (Williams, 2021). Subsequent analysis indicated a nationwide increase in racial tensions following Obama's victory in 2008 (Bigg, 2008). Some research has linked these spikes with the amplification of racial threat effects among the white ingroup of Americans (Wetts & Willer, 2018). The perceived shift of power from the ingroup to the outgroup is believed to have triggered perceived threats, which, for many individuals, justified their discriminatory behaviour

(Crandal & Eshleman, 2003). These observations are reflected in the sudden increase in racial tensions after the election with several key indicators (e.g., the KKK and Council of Conservative Citizens reported a significant rise in interest and membership in 2008; see Williams, 2021).

The links between racial tensions and Obama's election victory were also evident in the digital world, as there was a noticeable surge in hostile discourse and activity (Sela et al., 2012; Hankes & Zhang, 2017; Scrivens et al., 2021). When considering police-recorded hate crime data along with other indicators, such as increases in interest for far-right groups and spikes in online tensions, it is evident that Obama's election was associated with spikes in community tensions in the USA. Furthermore, many scholars view his administration as a turning point for the far-right movement (Simi, 2010), driven by the perceived threat of political and cultural change among the ingroup (Scrivens et al., 2021; Wetts & Willer, 2018). This thesis refers to casual factors associated with these upticks as "outgroup threat causes" in this thesis.

The few years after the 2008 election saw several instances of far-right-supported election wins, campaigning over issues of national identities, such as the election of Donald Trump and the Brexit referendum (Lubbers, 2019). Although some expectations that tensions would decline due to reduced perceived threats (Stephan & Stephan, 2002; Bigg, 2008), the opposite occurred. Scholars acknowledge that election victories and defeats can trigger tensions (Scrivens et al., 2021). Another marked increase followed Trump's 2016 election win (Perry et al., 2017; Edwards & Rushin, 2018; Müller & Schwarz, 2020; Bernatzky et al., 2022), with spikes larger than any point during Obama's eight-year administration (Edwards & Rushin, 2018).

Past studies have explored the concept of political legitimisation as a key explanation for the temporal causes of heightened community tensions (Perry, 2001; Koopmans & Olzak, 2004). Political legitimisation refers to validating bias attitudes by the state, party, or individual when elected, creating a hostile political environment (Perry, 2001). In light of this, Edwards and Rushin (2018) attributed the sudden increase in tensions after Trump's election to political legitimisation. They used the dangerous rhetoric theory to explain the origin and the validation theory to describe how political legitimisation enabled the acting of these attitudes (Edwards & Rushin,

2018). Thus, validation theory can be seen as an extension and facilitator of dangerous rhetoric theory. For example, although Trump made several inflammatory statements before winning the election, spikes in markers such as hate crimes were only visible after he was inaugurated, suggesting that once he was a successful candidate, many Americans perceived his rhetoric as legitimised and normalised (Okeowo, 2016; Reilly, 2016). This thesis refers to the political causes of tensions associated with legitimisation and validation theory as the "ingroup validation cause."

Bursztyn et al. (2017) tested validation theory's impact on divisive attitudes with a real-world example. In a study of 450 participants from states where Trump's win likelihood was 100%, two cohorts were formed: one informed of Trump's guaranteed win and a control group not informed. Results revealed higher anti-immigration attitudes in the informed group, aligning with Edwards and Rushin's (2018) validation theory. This theory posits that the perceived threat to the ingroup, accentuated by Trump's rhetoric, contributed to increased racial tensions, as he portrayed his supporters as economically strained and marginalised due to immigration (Edwards & Rushin, 2018). Both outgroup threat (e.g., Obama's election) and ingroup validation (e.g., Trump's election) are driven by the perceived threat from a marginalised outgroup and can be viewed as different means to the same end. They heighten threat and strain, which outweigh suppressions and justify prejudice (Crandal & Eshleman, 2003). Figure 4 illustrates these two politically driven causes of tensions that manifest into subsequent hate crimes and incidents, with the blue arrow symbolising outgroup threat and the red arrow symbolising ingroup validation. It's crucial to emphasise that the third stage (heightened sense of threat within the ingroup towards the outgroup) may manifest solely as increased threat and tensions often without progressing to hate crimes or incidents, as illustrated by the black arrow.

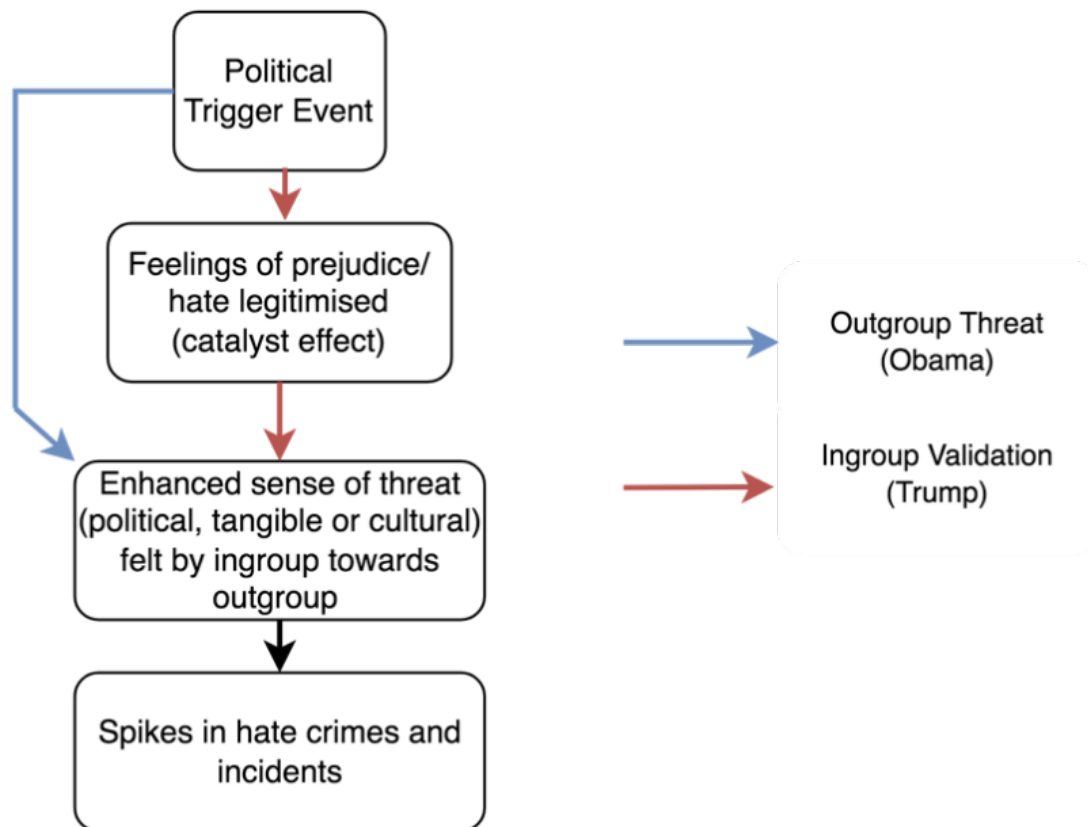


Figure 4: Political causes of tensions

Ingroup validation causes were not unique to Trump’s presidential election; another important example was evident in the UK during the Brexit referendum campaign and voted in 2016. A strong academic consensus shows clear correlations between the Brexit referendum and temporal shocks in tensions, that in many cases also boiled over into hate crimes and incidents (Moore & Ramsey, 2017; Piatkowska & Lantz, 2021; Devine; 2021; Williams et al., 2022). For example, there was a reported 23% of hate crimes in the direct aftermath concerning the previous average of 125 per day (Devine, 2018). These rises also controlled for other factors, such as terror attacks and the salience of immigration (Devine, 2018). Other evidence suggested that the overall prevalence of hate incidents was far greater than crimes recorded by the police. For example, initial evidence from Tell MAMA (2017) suggested a significant increase (+475%) in anti-Muslim street incidents was recorded over the weeks following the referendum in June 2016. These findings support previous notions that police-recorded data fails to consider the full extent of upticks in hate and community tensions (Nolan & Akiyama, 1999).

Moore and Ramsey (2017) investigated which central issues existed at the heart of the leave and remain campaigns. Findings suggested that the primary issue was immigration, closely followed by the economy (Moore & Ramsey, 2017). Although the economy was more widely discussed than immigration, subsequent analysis revealed that just under half (47%) of economic articles mentioned immigration, indicating that the two issues were frequently intertwined. Moore and Ramsey (2017) concluded that throughout the leave campaign, immigration was predominately covered negatively, with migrants painted as the source of economic and social problems by various social actors, including the media, activists, and politicians. These findings are consistent with Stephan et al.'s (2008) typology that perceived ingroup threat is caused by tangible (economic) or cultural (social) factors. After that, many economic arguments during the campaign were even regarded as a smokescreen for the underlying immigration issue and took overall precedence (Curtice, 2016). This negative coverage, predominately pushed by political parties, activists, and the media, fostered a divisive climate where immigration was perceived as threatening economic stability and cultural integrity (Burnett, 2017; Devine, 2018; Piatkowska & Lantz, 2016). This mirrors notions of dangerous rhetoric theory, exhibited in Trump's election Edwards and Rushin (2018). The upticks in community tensions associated with policy changes such as the Brexit referendum align with the dangerous rhetoric theory (King & Sutton, 2013) and similarly lend evidence to Edwards and Rushin's (2018) validation theory of political legitimisation. A notable distinction between the Brexit referendum and Trump's election is that significant upticks in markers such as hate crimes were observed during the campaign period before the actual vote (Devine, 2018). Although the increases were evident following the election, suggesting that the referendum result served as a political legitimiser of discriminatory attitudes, the uptick during the campaign can also be understood through the lens of validation theory.

Dangerous rhetoric during the campaign was heavily promoted by movements such as the Leave campaign and far-right political parties such as UKIP or BNP (Stewart & Mason, 2016; Modi, 2018), which lacked significant political legitimacy at the time. However, support for the Leave campaign and its associated rhetoric (i.e., immigration and the economy- (Moore & Ramsey, 2017) was not limited to unelected and extreme political entities. Many mainstream party members, across both left and

right affiliations, publicly supported Brexit (Hobolt, 2016; Vasilopoulou, 2016). Related literature often points to cases of extremism, such as terrorist attacks, as the trigger events most likely to cause upticks in community tensions (Legewie, 2013; Hanes & Machin, 2014; Williams & Burnap, 2016; Devine, 2018; Ozalp et al., 2020). While this has commonly been the case, with political events such as Obama's election being dwarfed by extremist (terrorist) events, some findings have indicated that political legitimisation during elections and votes can serve as a galvanising catalyst to enhance tensions. The political validation catalyst effect can be so significant that some examples have indicated that when present, they can outweigh those seen in extremist events (see Devine, 2018; Hopkins & Washington, 2020).

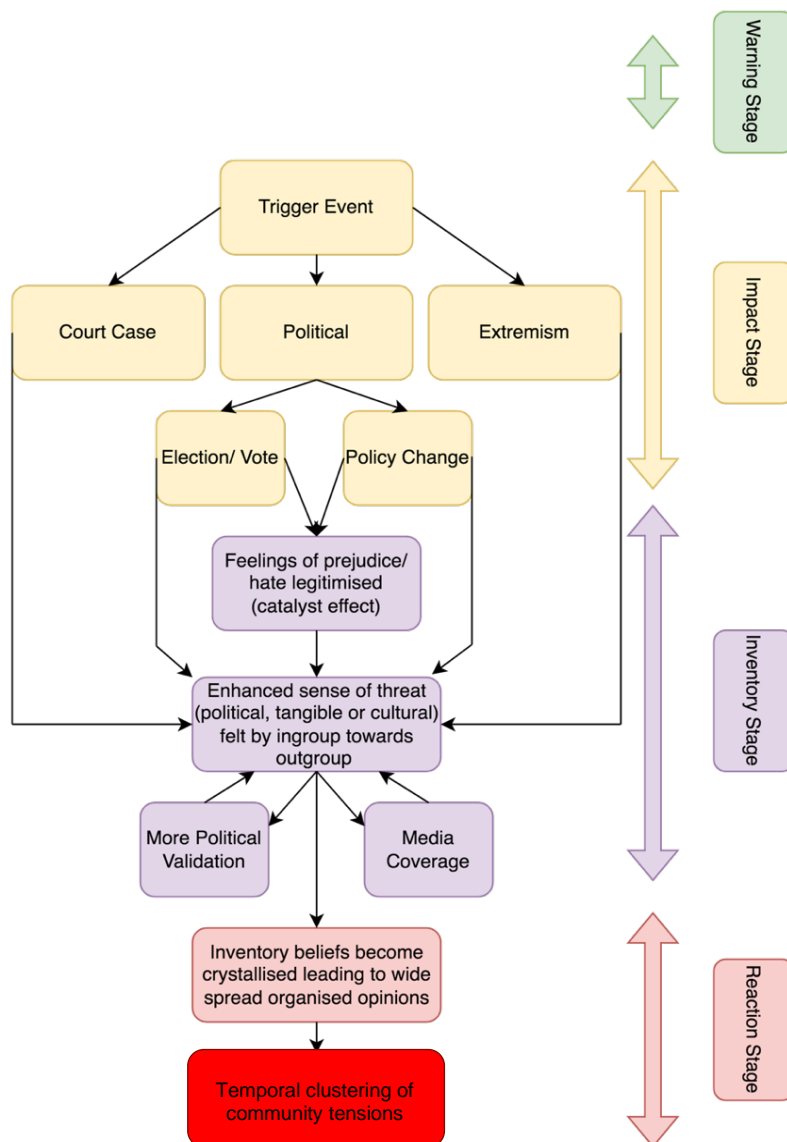


Figure 5: Literature consensus of trigger event typologies

### **2.5.5. Triggers and the Media**

Edwards and Rushin (2018) have made significant contributions to dangerous rhetoric by helpfully identifying political figures as key propagators of dangerous rhetoric. However, it is important to note that media sources have also been identified as influential in disseminating this type of rhetoric. Cohen (1972) developed a typology of deviant behaviour that comprises four stages: the warning stage, the impact stage, the inventory stage, and the reaction stage. This typology is commonly employed to explain the entire cycle of deviant behaviour and has been used to understand how the media can influence such events. Pre-existing prejudices and tensions characterise the warning stage before a trigger event. Media narratives and portrayals uphold and reinforce these prejudices (Knott et al., 2013). During the impact stage, the media takes on a more subdued role as they and the general public make sense of the event. However, it is towards the end of the impact stage and throughout most of the inventory stage that a pivotal time occurs when overall opinions relating to the trigger and broader social implications are formulated (Williams & Burnap, 2016). Cohen (2002) argues that this is a time when social actors shape the general public's opinions through agenda-driven narratives.

Moreover, empirical evidence has demonstrated the media's heavy involvement across all three typologies of trigger events (political events, extremist, and contentious court cases) (Hanes & Machin, 2014; King & Sutton, 2014; Moore & Ramsey, 2017; Sadique et al., 2018). For instance, studies have found that the USA's portrayal of Muslims in the media when discussing terrorist attacks has led to a spike in community tensions (King & Sutton, 2014). Other studies have shown that such media narratives exacerbate tensions during the aftermath of terror events and cause resurgences by keeping stories in the headlines (Sadique et al., 2018). Hanes and Manchin (2014) refer to these as attitudinal shocks reinforcing and intensifying intolerance and prejudice towards outgroups (another similar example relates to the Brexit referendum; see Moore and Ramsey, 2017).



### **2.5.6. *Micro-Triggers***

Given the considerable attention extreme triggers receive, the examination of micro-level triggers, which operate at local and regional levels, has understandably been overlooked owing to the nascent nature of this study area. Many note that discriminatory attitudes tend to manifest as a product of group behaviour rather than isolated occasions (Craig, 2002). While it is acknowledged that large-scale events significantly impact the prevalence of community tensions, this thesis posits that the group dynamic remains a crucial factor in localised trigger events and may even potentially lead to the emergence of more extreme behaviours. It is important to note that micro-triggers, as defined in this thesis, pertain to events that primarily involve local issues but can still have exogenous effects and attract national attention and news coverage. This is exemplified in Chapter Five, where the case of the Penally asylum accommodation illustrates the impact of such micro-triggers. Research on micro triggers is limited, and it remains uncertain if they have the same impact on individuals and communities as large-scale events. Most trigger event research delves into the effects on local and regional communities and relies on examining large-scale events for analysis (Craig, 2002; Legwie, 2013; Iwama, 2018; Albornoz et al., 2020; Awan & Zempi, 2020). There is a considerable need for further study of micro-level triggers, including those specific to local and regional contexts.

As noted earlier, most research has primarily focused on studying the spatial clustering of community tensions rather than temporal clustering (King & Sutton, 2013). This emphasis on spatial clustering provided important insights into the geographical distribution of community tensions. Still, it did not fully capture the dynamic nature of they unfold over time. That being said, considering spatial analyses helps indicate how trigger events may occur in local contexts. For example, one study examined how shifting demographics affect tensions in local neighbourhoods (Green et al., 2001). The study primarily aimed to investigate the spatial variations in community tensions, emphasising identifying the factors that may contribute to such variations. Factors such as the demographic composition, particularly the proportion of non-white individuals, and the perception of an area as a "white stronghold" are discussed as potentially influential determinants (Green et al., 2001). Garland & Chakraborti (2006) expanded on this by exploring rural

communities more generally, finding that negative discriminatory attitudes can emerge when such communities face events that disrupt their cultural homogeneity.

While these findings may not pertain to trigger events, as the demographic changes occurred relatively gradually (compared to other trigger events), they nonetheless suggest that the pace of causative factors of tensions can significantly impact localised areas. This highlights the importance of considering the temporal dimension in understanding the dynamics of community tensions in local contexts. The analysis presented in Chapter Five offers a unique opportunity to examine the effects of demographic change as a trigger event. The sudden influx of refugees in a short period due to a new policy announcing the establishment of asylum accommodation in Penally constitutes a micro-trigger event within a localised environment. Therefore, the analysis in Chapter Five contributes to the existing literature by providing the first examination of micro-trigger events where the event originates within the local area.

### ***2.5.7. Trigger Events and Digital Tensions***

The emergence of digital society has resulted in most facets of social life reorganising around online communications (Orton-Johnson & Prior, 2013; Shahidullah, 2018; Teresevičienė et al., 2018). Many initially praised this transition, arguing it would welcome a more connected, diverse, and accepting global society (Castells, 1996). Since then, this perspective has been challenged; although digital society has led to a more connected world, it has not become more tolerant (Bauman, 2013). Many point to the emergence of phenomena such as cyberhate as a key marker in heightened tensions, noting it has existed since the beginning of the public internet (Wall & Williams, 2007; Williams, 2006). Ample evidence has suggested that trigger events do not only lead to offline tensions but also spikes in digital ones, specifically antagonistic speech more generally that contribute to community tensions (Williams & Burnap, 2016; Ozalp et al., 2020; Czymara et al., 2022). In this thesis, 'antagonistic speech' is used to refer to speech detected online that is hostile, confrontational, or aimed at inciting conflict. Specifically, in the case of Chapter Five, this term is used to describe speech based on anti-refugee

sentiments. Moreover, considering this in the wake of trigger events is important from both a knowledge and methodological perspective for understanding digital tensions.

From a knowledge perspective, social communications migrating to digital spaces has presented several new issues when considering trigger events. First is the potential for increased ease of community tensions. Jaishankar's (2007) space transition theory suggests that the migration from offline to digital spaces can increase the likelihood of individuals expressing prejudiced feelings. This is facilitated by three main factors: identity flexibility, dissociative anonymity, and lack of deterrence (Jaishankar, 2007). The first two factors build on Festinger's (1952) deindividuation theory. Deindividuation refers to the process when individuals are more likely to engage in deviant activities or behaviours because they believe they cannot be personally identified (Festinger, 1952). As such, they undergo a process of disinhibition when they feel empowered to either offend or commit more serious offences than they may have otherwise (Joinson, 1998). As Jaishankar (2007) pointed out, social media and digital communications allow individuals to feel cloaked by anonymity, giving way to deindividuation and disinhibition. It is important to note that this process can take effect even when individuals are not anonymous. All this is required is perceived anonymity (Brown, 2018).

The final factor Jaishankar (2007) outlined is the lack of deterrence. Scholars note that phenomena such as cyberhate is heterogeneous and dynamic; 'it takes many different forms, and those forms can shift and expand over relatively short spaces of time' (Brown, 2018, p. 308). Combating it through policing or legislative frameworks is difficult (Bakalis, 2018). This means that deterrence is minimal in digital spaces compared to offline ones, especially when considering antagonistic speech more generally. Suppressive forces associated with being caught and identified are likewise very minimal. This is perhaps exacerbated in trigger events when more internet traffic is applied to an already stretched response system. Given that prejudice justifications are seemingly heightened during trigger events, the absence of significant suppressors can give away extreme upticks of digital tensions. The improvement of communications through technological advancements has enabled like-minded people to affirm and propagate each other's opinions (Earl &

Schussman, 2002). This builds on group socialisation theories before the emergence of digital society (Moreland and Levine, 1982; Levine and Moreland, 1994). Group socialisation refers to sentiments normalising in specific environments (Gallacher & Bright, 2021). When normalised, these sentiments and ideologies are subject to a sort of contagion effect in which they (i) spread to new people and (ii) radicalise those to whom they are exposed (Sunstein, 2017). This is hinged on the development of 'shared expectations and belief system about how to behave' (Gallacher & Bright, 2021, p. 3).

In a digital cohesion context, this has had many negative implications, with significant increases in digital tensions and so called "hate groups" on the internet (Perry & Olson, 2009). Therefore, technologies such as social media have been labelled as polarisation amplifiers that enable negative narratives directed at outgroups to be galvanised and propagated (Ellis, 2002; Daniels, 2008; Sunstein, 2017). Group socialisation and contagion theories are evident in online contexts, with a growing body of evidence showing that they contribute to the development and propagation of divisive content (Hirvonen, 2013; Oboler, 2016; Lima et al., 2018; Cinelli, 2021; Ziems et al., 2021). While contagions initially thrived on 'underground' platforms such as 8chan, 4chan, Gab and Voat (Lima et al., 2018; Williams, 2019), recent evidence has shown they occur on mainstream applications as well (Ziems et al., 2021)

The extent to which antagonistic sentiment online is socialised and spreads during shorter temporal bursts, such as trigger events, has also been examined (Williams & Burnap, 2016; Ozalp et al., 2020). These studies built on existing work by directly measuring the extent to which information or sentiment is "propagated by users, and which users have the most or least influence in the spread of such messages" (Williams & Burnap, 2016). Exploring information propagation of such factors is possible by including social media engagement metrics in subsequent analyses. The most frequently used metric for information dissemination on Twitter tends to be retweeted (Boyd et al., 2010; Williams & Burnap, 2016; Ozalp et al., 2020). During the analyses, information propagation (retweet count) and antagonistic speech were measured in two separate models. Predictors can be categorised into factor groups: content factors (i.e., hyperlinks, hashtags, sentiment), social factors (i.e., social actor

types) and external factors (i.e., Google searchers and news headlines) (Williams & Burnap, 2016; Ozalp et al., 2020). In addition, both studies control for the day of the week. In both studies, antagonistic speech was negatively correlated with information propagation (retweets). This indicates that although extremely high levels of online discriminatory attitudes were detected in the aftermath of trigger events, they received little public endorsement online. These suggest that it emanates from a small minority, “a core group who seek out each other’s messages... an ‘echo chamber’ of like-minded individuals who encourage and amplify each other” (Ozalp et al., 2020, p. 15).

Although neither study applied network analysis procedures, evidence of propagators operating in condensed groups resonates with the notion of core-periphery structures. Core-periphery structures were first identified by Friedmann (1967) and applied across a wide spectrum of academic disciplines (i.e., economics, biology, geography). Since then, they have been applied to social networks and structures in the social sciences (Borgatti & Everett, 2000). In this context, academics have debated what constitutes a core-periphery structure because many variations have emerged. To curtail such issues, Gallagher (2021, p. 1) provided an umbrella definition- “core-periphery structure is a fundamental network pattern, referring to the presence of two... distinct components: a dense ‘core’ of tightly connected nodes and a sparse ‘periphery’ of nodes loosely connected to the core and among each other”. Evidence in related work suggests that a core group publish antagonistic content, with most others not endorsing such beliefs (Williams & Burnap, 2016; Ozalp et al., 2020). As such, these findings are akin to core-periphery structures, with lead propagators existing in a condensed core and the non-endorsing public operating on the periphery. Analysis in Chapter Five expands on existing work by including network visualisations to demonstrate whether core-periphery structures of anti-refugee attitudes genuinely exist in cyber networks during trigger events.

The internet enables users to cluster according to shared interests rather than where they live (Castells, 1997). Individuals may interact and discuss particular events when and from where is most convenient. As such, wider audiences can interact with trigger events. This means the importance of spatial factors (i.e., the country where a

trigger event occurs) increasingly diminishes in digital spaces as such exogenous shocks can become more likely and are facilitated by the migration of social discourse to digital spaces. Exogenous processes predominately relate to information dissemination. Unfortunately, exogenous effects have also extended to antagonistic attitudes and even hate, labelled the 'globalisation of hate' (Schweppe and Walters, 2016:1). This has been corroborated in research in different contexts. For example, Williams and Burnap (2016) geo-tagged tweets after a terror attack in London, showing the shocks were reflected outside of both London and the UK<sup>6</sup>. Another example of digitally induced exogenous shocks is Trump's election in 2016, where analysis found that significant upticks of digital tensions and public interest occurred in areas outside the USA, including the UK and Canada (Perry et al., 2017; Giani & Méon, 2021). This thesis examines an extremely localised trigger event in a town in Wales. It builds on the surrounding literature by exploring whether local events can take on similar exogenous effects on information dissemination and tensions.

## 2.6. Summary

This chapter was divided into four sections, each dedicated to exploring distinct aspects of this thesis. Including a literature review and incorporating conceptual frameworks are integral components of rigorous and impactful research as they situate the research within the broader academic context and provide the necessary structure and theoretical underpinning for meaningful inquiry. The first section explored the emergence of digital society. This examination elucidated the digitisation of contemporary society and its profound impact on various facets of modern life. Next, the concept of community cohesion was scrutinised through the lens of the collective efficacy framework. Within these pages, a comprehensive understanding of cohesion was achieved, encompassing formal and informal community capacities. The third section investigated digital communication

---

<sup>6</sup> These results should, however, be interpreted with a degree of caution because approximately 1 per cent of Twitter users enable their tweets to be geo-tagged (Sloan et al., 2013). This presents a methodological dilemma for ascertaining truly reflective data of how online discussions spread according to user location in the aftermath.

strategies, employing Mergel's communication framework as the guiding framework. Preliminary insights were derived regarding how organisations harness digital technologies like social media to attain diverse objectives. Finally, the fourth section focused on the temporal aspects of community tensions, particularly about trigger events. Examining real-world case studies integrated with core conceptual frameworks and pertinent underlying theories (e.g., JSM and ITT) revealed how prejudice and community tensions could be initiated within digital and offline spaces. Concluding this chapter's comprehensive review, the following chapter will delve into the methodology underpinning this thesis, paving the way for a more profound investigation into the research's intricacies.

# Chapter 3

## Methodology

---

### 3.1. Introduction

This chapter provides an overview of the research methods and data used to investigate how stakeholders in Wales accomplish community cohesion. The approach is rooted in a mixed-methods pragmatic epistemology, emphasising the importance of using multiple methods to better understand complex social phenomena. The chapter is divided into five sections, each providing a detailed examination of the various methods employed. The start and end of this chapter provide a comprehensive overview of the mixed methods approach and ethical considerations. The remaining sections focus on specific study designs for three empirical chapters. The first uses primary survey data from Wales's "Cohesion Delivery Network" (CDN) to understand how stakeholders approach cohesion and tension issues. The second analyses the communication strategies of minority rights NGOs through social media data. The third explores digital tensions through social media data after a case study trigger event. These studies reveal how stakeholders in Wales address community tensions and the role of communication in shaping these dynamics.

### 3.2. Mixed Methods Approach

This research employed a mixed methods approach to provide a more comprehensive understanding of community cohesion in Wales. Given the multifaceted nature of community cohesion, various conceptual frameworks were considered, and corresponding methodological approaches were selected to provide a more comprehensive perspective. Mixed methods approaches have been widely used to examine multi-agency partnerships, with quantitative data often used to identify trends in collaboration within networks and qualitative data providing an



explanatory dimension as to why such trends exist (Gil-García, 2006; Olivet et al., 2010; Levi and Williams, 2013; Sedgwick et al., 2021). Scholars suggest that social media data lends itself to incorporating mixed methods approaches, as it provides rich textual and numerical data (Froehlich et al., 2020). Researchers have employed quantitative and qualitative practices in studying digital communication strategies. Qualitative analysis is often essential when conducting a content analysis of various communication strategies, as it helps identify and categorise these approaches thematically. Subsequently, quantitative methods are used to identify the overall prevalence of different strategies and determine the most predictive factors (Meijer & Thaaens, 2013; O'Connor, 2015). Recent studies have advocated using mixed methods research to comprehensively understand how and why organisations use different online communication strategies (Cullen, 2022).

Trigger event research has undergone significant methodological changes over the years. Initially, offline data sources were used, which combined quantitative inputs like police recorded data with qualitative inputs like interview data. However, in recent times, the trend has shifted towards emerging online data sources, such as social media data, which are more accessible and provide richer information. Despite this shift, mixed-method approaches have not declined. Instead, qualitative dimensions have become more nested in quantitative dimensions in trigger event research, primarily due to the advanced analytical procedures available to researchers from big social data (Edwards et al., 2013).

Furthermore, incorporating qualitative data in the analysis of communication strategies used by organizations is challenging, as individual actors are involved in trigger event communications. Hence, researchers often aggregate qualitative data and present it through word clouds alongside quantitative findings to comprehensively understand general themes (Williams et al., 2013). Such mixed-method approaches are commonly used in related research, and this thesis follows a similar methodology by utilizing social media sources and cross-sectional surveys

Defining "mixed methods" is a subject debated within the academic community. Generally speaking, it refers to the integration of quantitative and qualitative methods in a single study to gain a more comprehensive and profound understanding of a

phenomenon (Chen, 2006:75). Quantitative research is the collection of numerical data, which is analysed using mathematically based methods, such as statistics (Creswell, 1994:2). Qualitative research produces "findings not arrived at by exercising statistical procedures or other means of quantification" (Strauss & Corbin, 1990: 17), while quantitative research is objective and provides descriptive and inferential values (Creswell & Plano Clark, 2018). Integrating methods yields a more comprehensive understanding of the research topic due to the complementary strengths of each method (Morgan, 2013). In addition, qualitative research is inductive, and quantitative research is deductive (Morgan, 2013: 48). While scholars agree on these principles, the definition of mixed methods research remains contested. Patton (2002) defines mixed methods research as using different data sources and design elements to explore a question and bring different perspectives (p. 587). This approach involves collecting and analysing data separately to triangulate findings for an integrated final interpretation. However, some research has moved away from this notion. According to Schoonenboom and Johnson (2017), mixed methods research occurs whenever quantitative and qualitative components are combined at any point (p. 116).

The methodological literature suggests four potential points of integration (Morse, 2016): (i) the conceptualisation stage, where research questions and aims are formulated; (ii) the methodological experimental stage, covering the data collection period; (iii) the analytical stage, involving data analysis, and (iv) the inferential stage, concerning the interpretation of findings (Tashakkori & Teddlie, 2003; Morse, 2016; Schoonenboom and Johnson, 2017). Previous definitions of mixed methods research, like that of Patton (2002), have been restrictive by rejecting studies that integrate before the inferential stage. For instance, cross-sectional studies that collect both closed and open data were not considered mixed methods research since both forms of data were gathered concurrently. However, more recent definitions, such as those of Schoonenboom and Johnson (2017), support the view that both methods can and should be integrated well before the inferential stage, particularly when dealing with social media data, which is becoming increasingly integrated (Lomborg and Bechmann, 2014). Scholars note that social media content, comprises both quantitative and qualitative data, including the textual content of the post and its engagement levels (e.g., likes or views). Thus, researchers can take

advantage of the unique opportunities provided by social media data to collect both strands of data concurrently (Snelson, 2016).

Mixed methods approaches can vary, usually influenced by when and how the integration points occur (Schoonenboom & Johnson, 2017). Integration can occur concurrently or sequentially, depending on the order in which the methods are employed (Guest, 2013). If quantitative methods are used before qualitative methods in the analytical stage, this is known as explanatory sequential design, where the qualitative data is used to explain the initial quantitative findings (Creswell & Plano Clarks, 2011). Conversely, qualitative methods are used first in the exploratory sequential design, and quantitative data is subsequently employed to generalise the initial qualitative findings (Creswell & Plano Clarks, 2011). Multiphase designs involve two or more phases of either sequential or concurrent methods addressing a single research question (Creswell & Plano Clarks, 2011). Fully integrated mixed designs employ methods at every integration point, while multilevel mixed designs use concurrent methods at two or more integration points (Teddlie & Tashakkori, 2009). Convergent parallel mixed designs involve independent execution of both methods, and the results are brought together in the overall interpretation (Patton, 2002). In addition to the timing of methods, integrating mixed methods in research depends on the weighting given to each method. Both methods are sometimes given equal status, resulting in neither method dominating the final output (Doyle et al., 2009). However, when one method is dominant, the design is classified as either qualitative or quantitative dominant (Johnson et al., 2007: 124). These embedded designs have one dominant method, with the other data providing a secondary or supportive role (Caracelli & Greene, 1997).

### ***3.2.1. Epistemology/ Ontology***

Researchers are committed to their perceived version of the world, which shapes their approach to research questions in pursuit of legitimate knowledge (Hughes and Sharrock, 2016). These perspectives are known as the researcher epistemologies. It is crucial for researchers to be mindful of the framework that best aligns with their research question to produce cohesive research aims and design approaches

(Slevitch, 2011). The research paradigm debate centres around quantitative objectivism and qualitative subjectivism, each grounded in distinct philosophical approaches of positivism and interpretivism. However, this research adopts the pragmatic position, acknowledging the value of ontological and epistemological assumptions (Bryman, 2012). Pragmatism emphasises experience and its outcomes, rejecting traditional notions of reality (Morgan, 2013). This study deems the separation between objective and subjective epistemology unsuitable, and mixed methods present a third way that provides a comprehensive understanding of a single phenomenon. Pragmatism focuses on "what works" as the truth regarding research questions (Tashakkori & Teddlie, 2003: 713). It goes beyond integrating quantitative and qualitative methods in research and can be applied in all study phases. Scholars note that pragmatism is crucial in project design, data collection, data analysis, conclusions, and dissemination (Kelly and Cordeiro, 2020). For instance, data collection tool selection is an example of how pragmatism can be applied in a specific phase of the research process. Researchers can select data collection tools based on the required data components, not limited to quantitative or qualitative data but also the size of datasets and the variables they can produce. This is particularly significant when collecting social media data because some tools restrict collecting historical data and limit the amount of content that can be extracted. Therefore, researchers must adopt a pragmatist approach that focuses on "what works" to obtain the most appropriate data for the research question.

Pragmatist approaches are not simply the most convenient, as some academics argue. They offer flexibility and malleability, enabling researchers to capture the most useful data and analyse it tailored to meet the research question's requirements (Feilzer, 2010; Morgan, 2013; Teddlie & Tashakkori, 2003). Although pragmatism is often associated with mixed methods research, it has broader implications for research methodology. By incorporating a pragmatist approach in every stage of the research process, researchers can ensure that their methods align with their research questions and that their findings accurately reflect the reality they are studying. Throughout this chapter, specific examples illustrate how a pragmatic outlook interacts with each of the three empirical chapters.

### **3.3. Empirical Chapter 1**

#### **3.3.1. Context**

The first empirical chapter (Chapter Four) involved a network analysis of multi-agency partnerships responsible for delivering cohesion in Wales. In line with a pragmatist approach, a survey method was chosen to quantify different stakeholder groups and identify trends in communication patterns to answer the research question. The survey was distributed to seven key stakeholder groups, including the Welsh Government, police, cohesion coordinators, cohesion officers, community groups, and charities.

Unlike similar studies, the survey aimed to understand inter-agency partnerships between broad stakeholder groups rather than personal relationships. The study used multi-dimensional scaling (MDS) to visualise communication patterns between cluster groups, building on approaches previously used by Levi and Williams (2013). Analytical procedures, including MDS, principal component analysis (PCA), ordinary least squares regression, and qualitative data solicited from the survey, were used to achieve this. These methods provided comprehensive insights into a previously unexplored network. This section overviews Chapter Four's data collection procedures, pre-processing, and analysis methods.

#### **3.3.2. Data Collection**

The data in this chapter was generated from an online survey that included open and closed questions, allowing for the collection of quantitative and qualitative data (see Appendix 1 for the full survey). This was the first methodological integration point, as both data forms were collected concurrently (Bishop, 2014; Morse, 2016). Although collected concurrently, the methods were later analysed sequentially. With the advent of the digital age, online survey tools have become increasingly popular due to their cost-effectiveness, time efficiency, and potential to minimise biases that may exist in interview-administered surveys (Nayak & Narayan, 2019: 31). Despite the evident benefits of digital survey tools, scholars have criticised their use. According

to Nayak and Narayan (2019: 31), using self-completed digital surveys may result in poor sampling and incomplete or inaccurate responses.

In this study, there was no pre-established "cohesion delivery network" to draw on, as it does not exist in a formal capacity. Therefore, a randomised probability sample was not used as there was no defined population to target. Although non-probability samples have negative implications, Dorofeev and Grant (2006) suggest that these are minimised in studies that prioritise interrelationships between variables and use "soft measures" (Levi and Williams, 2013: 425). The survey data only indicates the nature of inter-relationships between stakeholder groups rather than overall prevalence. Using a non-probability approach ensured an equal balance of representation for node groups through selective targeting (Tabachnick and Fidell 2013).

According to Backor et al. (2007: 1), survey fatigue often results in participants abandoning the survey or submitting hurried and imprecise responses. Scholars have identified various factors that contribute to survey fatigue, such as an excessive number of questions, overuse of open-ended and demographic questions, and surveys that are too complex or appear unprofessional (Schonlau et al., 2002; Evans & Mathur, 2005; Sinickas, 2007). The pragmatic epistemological stance adopted influenced the survey design to minimise the risks of collecting inaccurate or invalid data. As a result, various guidelines were devised to ensure an effective survey design. In terms of survey length, the initial number of questions was reduced by Sinickas' (2007) suggestion of no more than 33 questions.

In addition, the survey design was informed through extensive consultations with board members and my supervisor. These discussions were crucial in identifying key themes, refining questions, and ensuring alignment with organisational goals. The board members' strategic insights and my supervisor's methodological expertise provided a robust foundation, addressing potential issues through iterative review and refinement. Given the delays caused by the pandemic, this comprehensive consultation process ensured the survey's clarity, relevance, and methodological soundness, obviating the need for a pilot study and enabling its timely implementation.

The final survey had 26 questions, including open-ended questions to capture qualitative data for explanatory insights into cooperation trends. Precautions were taken to prevent survey fatigue by including open-ended questions, primarily related to inputs relevant to the research question. All open-ended questions were optional with no minimum word count, and only essential demographic data such as stakeholder group and region were collected. The survey included three question types: unordered response categories, Likert-scales, and open-ended questions. To ensure clarity and accuracy, the survey focused solely on research-related questions, avoided technical or vague language, and refrained from asking multiple questions in one. Additionally, the survey's appearance and direction were made clear, and efforts were made to avoid long and leading questions. These criteria selected digital surveys as the most flexible and practical method for obtaining data to answer the research questions.

The survey had three dimensions. The first dimension (Q1-Q10) gathered participant-related information, including organisational details and perceptions regarding community cohesion. Questions were designed to elicit unordered response categories such as stakeholder group, region of operation, and the length of tenure in their current position. Additionally, two dichotomous questions were posed to determine funding sources. Likert-scales were used to measure self-perceived focus on tension detection, prevention, and response, as well as focus levels on the nine protected characteristics outlined in the Equality Act (Home Office, 2010). Finally, respondents were asked to predict the trajectory of tensions over the next five years using a two-question format.

The survey's second dimension (Q11-Q18) explored how respondents perceived cooperation within the cohesion delivery network. The initial questions focused on the importance of cooperation and opportunities for improvement. Using a five-point Likert-scale, respondents were asked to rate the current frequency of cooperation between their stakeholder group and each of the cluster groups. They were also asked whether their group should cooperate more, less, or to the same extent with each cluster group. Additionally, an open-ended question allowed respondents to elaborate on their current level of cooperation frequency and reasons for wanting to alter it.

A Likert-scale matrix with five points was used to measure stakeholder group cooperation quality, and an open-ended question was added for responders to explain their scores. The survey also included questions (Q19-Q23) about the effects of COVID-19 on the network and social media use in operations. Two five-point Likert-scales were used to gather responses on the impact of COVID-19 on work and cooperation with other agencies, followed by open-ended inquiries. Finally, respondents were asked to rate their belief in implementing more social media monitoring and their organisation's use of social media for inclusive content on a five-point Likert-scale.

### ***3.3.3. Survey Dissemination and Ethical Considerations***

Qualtrics proved an effective and cost-efficient tool for collecting data from CDN stakeholders, yielding n=59 complete responses without any partial or incomplete submissions. Respondents were targeted through professional channels to prevent the survey from being overlooked in junk mail, and the design and communication were professional. The survey flow was guided using Qualtrics' help materials to minimise issues.

The survey questionnaire included an information sheet and a consent sheet. The information sheet provided a comprehensive overview of the study and the kind of data being collected, while the consent sheet allowed participants to opt in or out of the study after reading the information sheet (refer to Appendices 2 and 3). To address ethical concerns associated with obtaining consent from digital survey participants, an e-signature function was integrated at the beginning of the survey. This approach aligns with the ethical practices outlined by scholars who have identified potential issues with obtaining signed or vocal consent from participants using digital surveys (Regmi et al., 2016).



### **3.3.4. Preprocessing**

NVivo and SPSS were used to process the data for analysis. NVivo facilitated thematic analysis of textual data, while SPSS allowed for quantitative analyses, including MDS, PCA, and regression models. In SPSS, the data was cleaned, and variable names were changed to enhance navigability and exploration. The initial step was to identify potential missing values. Diagnostics revealed that none of the closed items were partially or entirely incomplete. As the survey included ordinal and categorical data instead of scale variables, additional steps were taken to prepare the data for analysis. Categorical variables were recoded into numerical values or dummy variables, enabling an appropriate data format for analysis.

During data preprocessing, open-ended survey responses were prepared for analysis in NVivo. The data was organised into nodes to categorise and group relevant themes or topics related to the research question (Basit, 2003). Themes and topics were manually identified and assigned to relevant segments discussing each of the seven stakeholder groups. The coding framework consisted of two core elements: identifying stakeholder groups and categorising thematic motifs. The coding process for the former adhered to a systematic protocol that thoroughly examined each response. This process encompassed the allocation of authorship and subject stakeholder group codes, supplemented by cross-referencing techniques when deemed necessary. The iterative nature of this approach facilitated a comprehensive and nuanced analysis, thus establishing the foundational framework for subsequent analysis. The latter codes were utilised to stratify responses into overarching themes, such as "social media impact on the network," "cooperation patterns (quantity) and (quality)", and Covid impact aiming to discern recurrent patterns. This methodically structured coding paradigm unveiled stakeholder perspectives within community cohesion initiatives. Finally, the data was cleaned to remove irrelevant information and errors, ensuring it was ready for analysis.

### **3.3.5. Methods of Analysis**

As previously mentioned, quantitative and qualitative data integration commenced during the data collection phase, where both data forms were gathered through open and closed fields within an online survey (Morse, 2016). Although collected concurrently, the data sets were later separated and analysed in specialised software programs according to their respective data types during the preprocessing stage, resulting in a more comprehensive understanding of both data forms.

Data separation persisted into the analytical phase, with quantitative analysis preceding qualitative analysis. An explanatory sequential design was used (Bowen et al., 2017), where quantitative analysis revealed patterns in the cohesion delivery network among stakeholders, and qualitative data provided insights into underlying reasons. Four forms of quantitative analysis were used: initial univariate and bivariate descriptive analyses, MDS, PCA, and ordinary least square regression models. Thematic analysis explored the qualitative analysis, examining open data extracts (Braun et al., 2021). The following section will elaborate on the methodological procedures for each approach.

### **3.3.6. Multi-Dimensional Scaling (MDS)**

During the analysis phase, survey data was input into a series of MDS models. MDS, a widely adopted methodological approach, serves to discern patterns and relationships within data. This method operates by utilising sets of quantitative proximity data, which may comprise distances or similarities between objects or nodes and represent them visually as space points. By doing so, complex numerical data such as quantitative matrixes can be more readily interpreted intuitively (Coxton, 1982). In this vein, geometric distances between nodes in visualisations can indicate empirical relationships in the data, offering a comprehensive understanding of its underlying structure. Originally, MDS was used to visualise geographic proximities based on metric measurements, such as city distances in miles, plotted on two axes - North-South and East-West - to represent coordinates (Kruskal and Wish, 1978). MDS has since been used in non-geographic fields like

marketing, mapping similarities and dissimilarities between chocolate bar brands using Likert-scale ratings (Janssens & Wijnen, 2012).

MDS has been gaining popularity in social sciences, particularly in criminology, for offender profiling (Salfati, 2000) and is now being used in network analyses (Naurin & Lindahl, 2007; Levi and Williams, 2013) to map cooperation patterns between stakeholders. Unlike geographical proximities, MDS in social sciences uses non-metric measurements based on perceptions of similarities or dissimilarities between objects, often gathered through ordinal data (Van der Maaten & Hinton, 2012). This provides a unique opportunity for researchers to uncover underlying themes in cooperation patterns and networks. Despite validity issues, collecting non-metric data can be useful, particularly when individual perspectives are central to the study because they are inherently subjective and relate to human perceptions (Levi and Williams, 2013: 425).

MDS was implemented three times, with the first two measuring cooperation patterns among key stakeholders (Naurin and Lindahl, 2007; Levi and Williams, 2013), following procedures from previous studies. The first application used Likert-scale items to gather information on cooperation frequency. Time indicators corresponding to response options, such as "several times a week" or "weekly", were introduced to reduce potential data validity issues due to subjectivity. The second item was more subjective, asking respondents to rate the quality of their cooperation with other stakeholders, so a basic direction of scale was used (e.g., 1 poor quality and 6 excellent quality). These two items were applied to seven nodes representing stakeholder groups, and the PROXSCAL process in SPSS mapped cooperation frequency and quality into an MDS space. The final MDS application focused on the protected characteristics identified by the Home Office (2010) using a directional scale and the PROXSCAL procedure, with proximity indicating clustering of focus between nodes.

### *3.3.6.1. Factor Analysis*

Factor analyses were conducted to reduce the number of variables into related meta-clusters with fundamental similarities based on measuring the inter-correlations between stakeholder groups. The factor analysis was performed using the same survey items included in each stakeholder group's MDS procedure, with PCA and varimax orthogonal rotation. In this context, the MDS findings reflect the PCA factor loadings and offer a visualisation and additional dimensions. This process aimed to identify a smaller number of meta-clusters that capture the inter-correlations of stakeholder groups, thereby reducing the data for later analysis. The resulting components produced by the PCA are labelled as meta-clusters, resembling stakeholder group clusterings<sup>7</sup>.

### *3.3.6.2. Ordinary Least Square Regression*

Whilst MDS procedures offer valuable quantitatively derived visualisations of the CDN cooperation space, their effectiveness in determining other factors influencing cooperation is limited. As a result, researchers have turned to regression models to gain additional insights into the factors that underlie cooperation in action networks, as Levi and Williams (2013) have demonstrated. By employing this methodological approach, scholars have gained a more comprehensive understanding of the various factors contributing to effective cooperation in such networks. After the PCA, three Ordinary Least Squares (OLS) models were developed, each corresponding to the identified meta clusters. However, due to limited sample size and violation of the normality assumption, the BCa bootstrapping method was used to estimate the statistics' sampling distribution<sup>8</sup>. Certain predictors were removed due to multicollinearity issues, but correlation and tolerance statistics showed a robust fit for all models after removal.

---

<sup>7</sup> Further information is detailed in the results (Chapter Four)

<sup>8</sup> This method adjusts the confidence interval to account for inherent biases and asymmetries (Haukoos and Lewis, 2005).

The three component meta-clusters produced in the cooperation frequency factor reduction were transformed into separate scale variables (Table 9). These scales were used as dependent variables in three separate models. The models, therefore, related to the cooperation frequency observed with each of the three discovered meta-clusters. The three models had ten predictors categorised into two sub-factors: (1) cluster characteristics and (2) cluster perceptions. The former had six predictors: high-level detection, prevention, response, multi-focus and region, EU transition funding, and government-based. Multi-focus was a dummy variable created by coding respondents with a high-level focus on two or more meta-cluster groups. The latter sub-factor aimed to understand the perceived cooperation quality for each of the three meta-clusters. The three meta-clusters were reduced using PCA in relation to the quality of cooperation Likert-scale items considered in the MDS procedure. Sub-model analysis showed which predictor set explained the most variance in cooperation frequencies for each model.

### **3.4. Empirical Chapter 2**

In recent years, scholars have increasingly investigated the temporal aspects of community tensions rather than just its spatial dimensions (Williams, 2021). The concept of "trigger events" has been used to describe the causes of temporal shocks in tensions and associated behaviours. In the nascent field of trigger event research, new analysis methods have progressed through multiple iterations. Initial studies relied on offline data sources, such as interviews, police records, and surveys (Legewie, 2013; Hanes & Machin, 2014; King & Sutton, 2013). While offline sources were considered robust and the "gold standard" for vigorous sampling procedures, they were also criticised. Williams and Burnap (2015: 214) highlight three limitations of using offline data: 1) low temporal granularity, 2) retrospective reporting and recall biases among witnesses and victims, and 3) underreporting in official police data, particularly in the context of hate crimes. Consequently, scholars believe that many offline data sources fail to capture the full extent of community tensions and their impact (Ozalp et al., 2020: 3).

Social scientists and criminologists now suggest using big data to overcome the limitations of offline data. Social media data is utilised to obtain real or near-real-time insights into the expression of genuine emotions due to its timestamp features (Edwards et al., 2013; Manjubashini et al., 2022). This method marks a significant shift in trigger event literature and expands the interdisciplinary approach to computational criminology by covertly recording and analysing post-trigger event information flows (Drouhot et al., 2023). Because online trigger event research is an emerging and evolving focus area, it has led to a fragmented approach within the academic community. Researchers often develop in-house classifiers, resulting in variations among these methods. This chapter proposes a more standardised and accessible alternative to using in-house machine learning classifiers, as proposed by most digital tension trigger event studies except for Czymara et al. (2022). The proposed method uses the Google Perspective API to measure antagonistic speech during trigger events. This method is quick, easy, free to use, and replicable in future studies, making it suitable for researchers who lack access to computer scientists. However, using "one shoe fits all" classifiers is not without criticism, as will be explored later in this section (Kumar et al., 2021: 299). This section will delineate the five distinct methodological phases utilised to examine the trigger event case study concerning the announcement of asylum accommodation in Penally, including the use of the Google Perspective API.

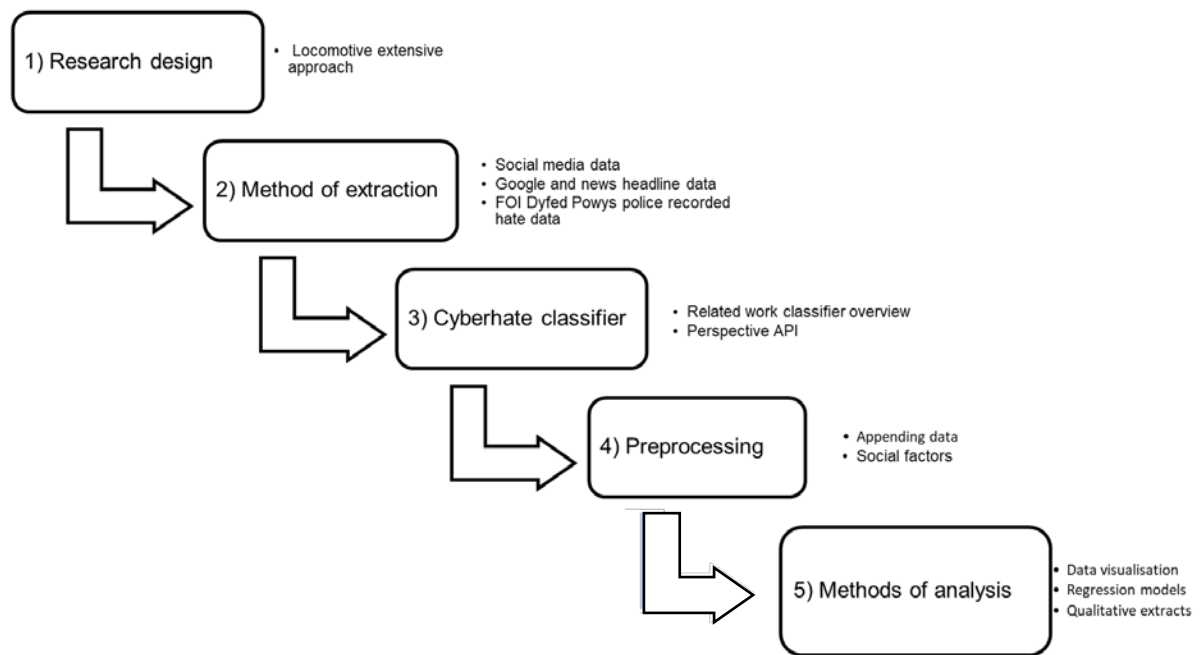


Figure 6: Methodological steps taken in Chapter Six

### 3.4.1. Research Design

Scholars propose four research strategies for analysing interactions through social media data: locomotive intensive, locomotive extensive, punctiform intensive, and punctiform extensive (see Table 1). These strategies reflect conventional research designs used before the emergence of digital society and are now applied in relatively new domains such as digital tension research (Edwards et al., 2013; Williams & Sloan, 2015). Locomotive studies observe social relations as a dynamic process, while punctiform studies examine snapshots of these relations (Sayer, 1992). Traditional punctiform methods that focus on the location of tensions rather than the time of occurrence have significant design flaws when attempting to apply these methods within a temporal framework. Williams and Burnap (2015) argue that low temporal resolution, retrospective reporting, recall biases among witnesses and victims, and underreporting in official police data can lead to findings that lack validity, particularly in the context of community tensions. As the academic focus has gravitated towards digital tensions and temporal dimensions, an increasing number of studies have adopted locomotive approaches (Burnap et al., 2015; Williams & Burnap, 2015; Williams et al., 2019; Ozalp et al., 2020; Czymara et al., 2022). The

methodological decisions are based on a pragmatic approach that considers the demonstrated success of locomotive approaches in contrast to the limitations of previous punctiform methods (Williams & Burnap, 2015).

*Table 1: Design strategies for analysing social relations with examples (Edwards et al., 2013)*

		<i>Research Design/ Data</i>	
		Locomotive	Punctiform
<i>Research Strategy</i>	Intensive	E.g., Ethnography/ Observational studies	E.g., Cross-sectional qualitative interviewing
	Extensive	NSM Analysis (capturing naturally occurring data in real/useful time at the population level)	E.g., Surveys (cross sectional, Longitudinal) and experimental studies

Big data available on social media allows researchers to conduct extensive locomotive assessments of real-time trigger events (Manjubashini et al., 2022). This approach provides insights from a wider range of social actors than traditional survey and interview methods, allowing for a fine-grained analysis of phenomena that can be examined to the second (Edwards et al., 2013). For example, perpetrator are unlikely to participate in punctiform exercises like interviews or surveys, making extensive locomotive approaches more reliable for establishing realistic perceptions and attitudes in real or near-real time (Williams & Sloan, 2015). Covert approaches can mitigate the risk of bias (Drouhot et al., 2023). Data extraction methods were subsequently considered in light of the preferred use of extensive locomotive approaches.

### **3.4.2. Method of Extraction**

An eight-month data collection period from August 2020 to March 2021 aimed to measure the announcement, duration, and closure of the Penally asylum accommodation and associated online activities. Triangulating extensive locomotive social media data with curated and administrative data can enhance the richness and scope of data and result in more comprehensive analyses (Edwards et al., 2013; Sloan et al., 2013). A triangulated approach was adopted using data from four



sources, reflecting a preference for data augmentation over big data surrogacy (Edwards et al., 2013). However, the relative importance of the different data sources used varied, with some dominating the final output and others contributing only marginally to specific findings, as explored later in this section.

The primary data source was derived from Twitter. Social networking platforms were considered the most suitable option for gaining insights into online opinions and discussions involving diverse social actors, and the selection process saw Twitter chosen over other platforms due to factors such as the ease of its API data extraction process, balanced modalities, and non-reciprocal structures in public spaces (see Appendix 4). In trigger event studies, open-networking structures are essential. Open friendship networks allow all tweets to be publicly viewed without requiring followers to follow the author. They create a digital version of a public square or agora (Williams & Burnap, 2016: 218), enabling citizens to discuss trigger events openly. This is further facilitated by platform functions like trending topics and hashtags, which centralise discourse around any given event. Twitter was chosen as the primary platform for data extraction based on a pragmatic approach, where the "public square" nature of the platform was deemed most effective in addressing the research questions.

After selecting Academic Twitter<sup>9</sup>, the extraction diagnostics considered the collection period and keyword. For the Penally asylum accommodation study, data was collected between August 1st, 2020, and March 31st, 2021, using a keyword search for a reliable and non-trivial sample (Yang & Counts, 2010). Keyword search terms are particularly useful in information diffusion and trigger event research. Williams and Burnap (2015) used the highest trending word, "Woolwich", in a two-week window following Drummer Lee Rigby's murder, finding that the collected tweets using the "Woolwich" search term were mainly related to the trigger event.

After manually inspecting the highest trending keywords, 'Penally' was identified as the most prominent term, generating significant online activity related to the asylum accommodation, as expected for smaller-scale triggers with limited pre-trigger online

---

<sup>9</sup> A more detailed explanation for Academic Twitter's selection is provided in Appendix 11.

traffic. Therefore, a census of all Twitter activity (N=6,661) related to the 'trigger event' was collected using the 'Penally' keyword between 1st August 2020 and 31st March 2021 through Academic Twitter, allowing concurrent quantitative and qualitative data collection. Big data was the primary source among the four used in this study, supplementing traditional data sources and expanding research designs not typically applied in extensive locomotive contexts (Williams & Sloan, 2015; Morse, 2016).

### ***3.4.3. External Factor Data Sources***

This study's second and third data sources are external factors included in the regression models, which aim to assess how other forms of open data interact with online activity, building on previous research (Williams and Sloan, 2015). The first external data source measures the impact of media coverage by analysing news headlines published daily over eight months. This metric aims to determine the influence of parts of the media on overall online discourse regarding the asylum accommodation. News contributions from social media platforms were excluded, as other predictors account for them (see social factors section).

Various tools may be used to extract headline data, such as LexisNexis, NewsBank, and ProQuest, as identified by Weaver and Bimber (2008). Free tools like Slashdot, Google News, and AOL News are also available, which are more accessible. Weaver and Bimber (2008) found that Google News captured more news stories than LexisNexis. However, an exploratory analysis revealed that many headlines included blog entries, which do not qualify as "traditional news" in this study. While some studies may find blog entries noteworthy, they were of less concern to this research. Each of the three traditional news data collection tools can exclude non-traditional news sources and collect local, national, and international news.

The study's focus on a local trigger event necessitated prioritising local news outlets. Weaver and Bimber (2008) suggest LexisNexis is the most efficient news archive, and its accessibility and capacity for localised collections further motivated its selection. LexisNexis permits researchers to customise searches using criteria such

as keywords, time spans, and outlets. The pragmatic selection of LexisNexis<sup>10</sup> as the optimal tool to address the research questions was guided by factors such as traditional news filters and researcher accessibility. To collect news articles about the Penally asylum accommodation, the search term "Penally" was entered into LexisNexis, omitting new media sources (i.e., blog posts). However, many collected articles were off-topic, discussing tourism in Pembrokeshire instead of the asylum accommodation. Therefore, the search terms were refined to "Penally" and "Camps" to eliminate irrelevant cases. Although this refinement carried the risk of excluding relevant articles lacking the keyword "camps," an inspection of the updated dataset found no remaining off-topic cases. The resulting news dataset from media outlets displayed the daily frequency of headlines containing the keywords "Penally" and "Camps," comprising 550 headlines with an average of 2.26 per day (range 0-20) during the study period.

The Google Trends service was used as a third open data source in the study to examine how online interest in a trigger event can impact digital activity. Google was preferred over other search engines due to its widespread usage in Western countries (Purcell et al., 2012). Search terms entered into search engines can be correlated with offline and online phenomena, as demonstrated by previous research linking flu-related search terms on Google to visits to local doctors (Ginsberg et al., 2009). Digital trigger event studies have also shown a relationship between Google searches and the volume of information propagated on social media (Williams and Burnap, 2016). However, no significant associations were found with antagonistic speech, indicating that news headlines may have a greater role in exacerbating online tensions than search engines. Nevertheless, non-social media information flows can still impact events in digital spaces and should be included in the analysis. However, confirming that search engine data was related to the asylum accommodation trigger and non-trivial was challenging. Using a pragmatist approach, the keyword selection that yielded more accurate results in the news headline collection ("Penally" and "Camps") was preferred. The resulting dataset

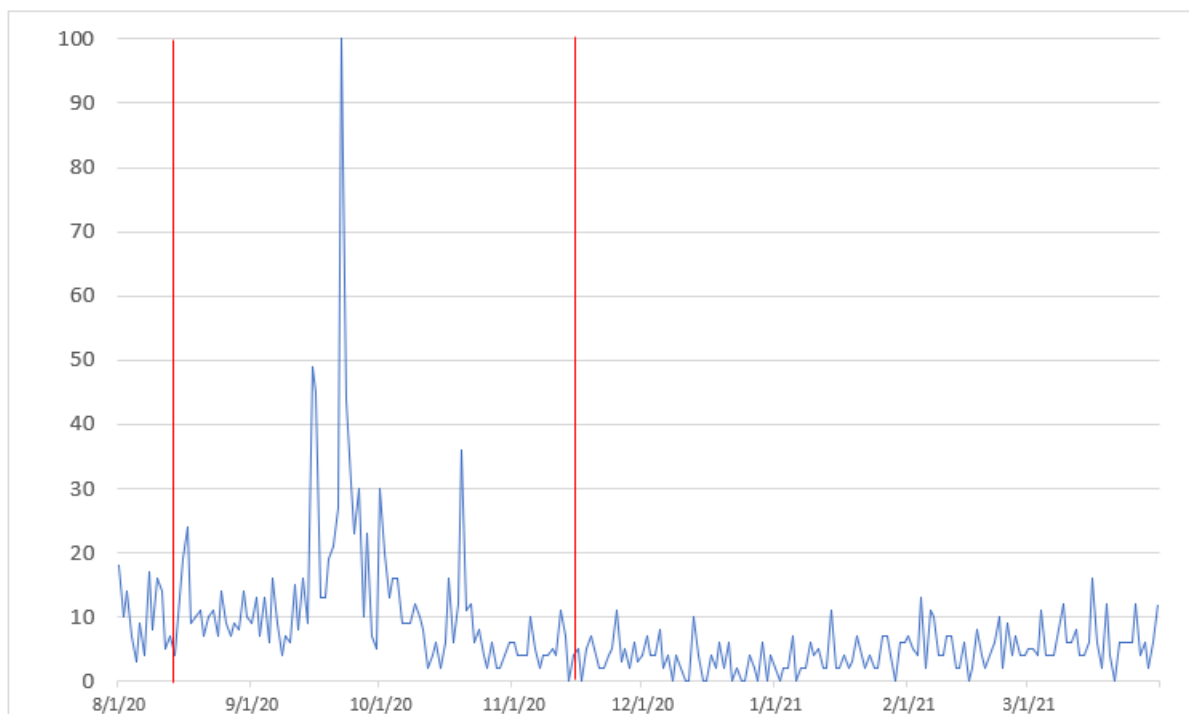
---

<sup>10</sup> LexisNexis is the only "traditional" news database tool freely available at Cardiff University.

indicated levels of the specified search terms per day as a percentage of the highest-ranking day and was saved as an Excel file for the preprocessing phase.

#### **3.4.4. Police Recorded Hate Incident Data**

The last data source concerns the controls utilised in regression models. Previous research suggests that time factors are crucial influences on information propagation flows on social media (Zarella, 2009; Williams & Burnap, 2016). However, most past studies focus on large-scale, short-term trigger events, limiting the time factors to time of day and day of the week. Williams and Burnap (2016) used Google search engine data to derive an issue attention cycle in the aftermath of the Woolwich terror attack, lasting 15 days. This cycle was mapped out over a condensed period, indicating the escalation, duration, diffusion, and de-escalation of online tensions and information flows. Following this approach, and in line with the pragmatic approach taken in this study, Google search engine data was used to establish an issue attention cycle:



*Figure 7: Issue attention cycle, based on Google search data*

Following this approach, Google search engine data was used to establish an issue attention cycle for Penally, which lasted 90 days. Police recorded hate crimes, and incident data were also collected using an FOI request. They were used to establish the months with the highest hate crimes and incidents, helping to create a reference category which can be found in Table 2.

*Table 2: FOI figures for hate crime and incident (Dyfed-Powys Police)*

<i>Month</i>	<i>Hate Crimes</i>	<i>Hate Incidents</i>	<i>Hate (Total)</i>
<i>August 20</i>	0	0	0
<i>September 20<sup>1</sup></i>	21	4	25
<i>October 20</i>	11	4	15
<i>November 20</i>	13	1	14
<i>December 20</i>	5	0	5
<i>January 21</i>	7	1	8
<i>February 21</i>	4	0	5
<i>March 21</i>	6	0	6
<i>Total</i>	67	10	77

With September exhibiting the highest levels of hate encounters (both incidents and crimes), it was selected as the reference category in the regression models (see methods of analysis section). With the FOI request data saved in an Excel file, the four separate data sources were all collected -(1) Academic Twitter data, (2) LexisNexis news headline data, (3) Google search engine data and (4) FOI Dyfed-Powys Police hate encounter data.

### **3.4.5. Anti-refugee Content Classifier**

#### *3.4.5.1. Classifiers in Research*

Social media data has gained popularity in studying public sentiment trends following trigger events, offering advantages over traditional offline data sources (Bernatzky et al., 2020). However, since studying trigger events in relation to digital tensions is a new research area, scholars have had to use interdisciplinary methodologies to develop in-house machine classifiers for identifying antagonistic, and sometimes even hate speech on social media platforms (Burnap et al., 2015; Alorainy et al., 2019; Williams et al., 2019; Ozalp et al., 2020; Manjubashini et al., 2022). Though these classifiers show promise, their use is limited. In-house classifiers are not

widely available, which leads to unnecessary intellectual duplications or barriers to entry for researchers without the necessary skills or access to computer scientists (Salminen et al., 2020).

#### *3.4.5.2. Perspective API*

Developing bespoke machine classifiers to analyse toxic or antagonistic speech on social media platforms is challenging, resulting in a growing need for accessible, general classifiers. Collaborations between disciplines are necessary, but the availability of public classifiers can prevent needless duplication of effort and remove barriers to entry for researchers. There is growing interest in standardised classifier tools, such as Bag of Words and Word2Vec to address this need. Preliminary studies have evaluated these tools' effectiveness in detecting antagonistic speech from textual data, with Google's Perspective API<sup>11</sup> reportedly being the most accurate among them (Mansourifar et al., 2022: 13). Thus, the Perspective API was chosen as the anti-refugee classifier in this thesis.

Computational social science tools like Commanalytic have allowed scholars to use Perspective API classifiers on social media data sets with multiple text string cases (Gruzd & Mai, 2021). Scholars now use perspective API to measure bias, antagonistic and even hateful content on social media (Sahi et al., 2018; Czymara et al., 2022; Mansourifar et al., 2022; Salehabadi et al., 2022). It uses machine learning algorithms to analyse various emotional concepts and provides scores from 0 to 1, reflecting the probability that a reader would perceive the comment as containing a given attribute (Google 2021b). The six attributes<sup>12</sup> include identity attack, toxicity, severe toxicity, profanity, insult, and threat, with definitions provided by Google (2021a) and summarised in Table 3. Google recommends using a threshold of 0.7 to indicate the presence of an attribute.

---

<sup>11</sup> Perspective API was created in February 2017 by Jigsaw and Google to help web publishers identify and moderate comments that undermine civil discourse (Google, 2022).

<sup>12</sup> Since this writing, Google has expanded the list of attributes to include sexually explicit, obscene, and flirtation (Abburri et al., 2021; 406).

Table 3: Perspective API classifiers cases Google (2021a)

<i>Definition</i>	
<i>Identity Attack</i>	Negative post attacking someone because of their identity (including race, gender, sexual orientation, ideology, religion, nationality, etc.)
<i>Toxicity</i>	Rude, disrespectful, or unreasonable post
<i>Severe Toxicity</i>	Very hateful, aggressive, disrespectful post. This score is less sensitive to posts that include positive uses of curse words
<i>Profanity</i>	Post with swear words or other obscene language
<i>Insult</i>	Insulting, inflammatory, or negative post toward an individual or a group
<i>Threat</i>	Post with an intention to inflict pain, injury, or violence against an individual or group

Researchers use identity attack to measure antagonistic speech related to factors such as religion, race, sexuality, and gender, while toxicity is used to measure other social tensions like general trolling and ideological conflicts (Alorainy et al., 2019; Czymara et al., 2022; Mansourifar et al., 2022; Strathern & Pfeffer, 2022).

Perspective API has been increasingly used to study digital tensions and cyberhate; however, its application has primarily been applied more generally rather than specifically temporal shocks like trigger events. Czymara et al. (2022) used Perspective API to measure temporal shocks in hate following terror events in Europe. This was limited to contentious videos with anti-immigrant tones and did not explore wider trends in online discourse or generalisable shifts in perceptions resulting from trigger events. This thesis presents a methodological innovation in applying Perspective API to trigger events or temporal shocks in anti-refugee content towards generalised discourse online, offering a more comprehensive understanding of how such events can influence the flow of information and the occurrence of digital tensions. This allows for the first application of Perspective API to the analysis

of a wider range of social actors, including politicians, activists, the media<sup>13</sup>, and police, as well as general citizens, rather than just those who comment on contentious YouTube videos.

Publicly available classifiers like Perspective API reduce the need for researchers to repeatedly develop similar tools, making digital tension research more accessible (Salminen et al., 2020: 2). However, Perspective API has limitations. Scholars criticise it for being too general and failing to identify offensive terms towards specific minority groups, making it less reliable than hand-coded approaches (Kumar et al., 2021: 299). The Perspective API faces challenges in comprehending language within particular contexts, leading to notably lower offensive scores. Introducing a full stop or including the word "not" can alter the context of statements and impact the scores (Hosseini et al., 2017). These double negatives contribute to false alarms.

For this study, a census of all tweets in the dataset was processed using perspective API on the Communalytic tool (see Appendix 5 for overview of process). The final output included all six available attribute scores (toxicity, severe toxicity, insult, threat, profanity, and identity attack) and one sentiment score. When assessing which Perspective API attribute to use as a measure for anti-refugee attitudes, the API definitions were considered, with 'identity attack' being deemed the most appropriate because it specifically detects language that negatively targets minority groups relating to the protected characteristics (Home Office, 2010). Before confirming whether or not to use identity attack, several measures were taken to ensure it would generate results with adequate validity. First, the researcher evaluated the accuracy of the identity attack in relation to the data. For classifiers such as Perspective API scholars often employ confusion matrices to calculate precision, recall, and F1 coefficients<sup>14</sup> by employing true positive, true negative, false positive, and false negative cases (Mansourifar et al., 2022).

---

<sup>13</sup> The media was predominately measured by exploring written news headlines and readers should note other media forms could exhibit notable differences.

<sup>14</sup> Recall = TP/ (TP+ FN), Precision = TP/ (TP+ FP), F1 = (2\* Recall \* Precision) / (Recall+ Precision) (Note: T = true, F = false, P= positive and N = negative)



The API scale was converted to binary data, with values above 0.7 categorised as an 'identity attack' based on Perspective's guidelines (Google, 2021b). To assess classifier accuracy, 500 tweets were hand-coded by two researchers (Dhaoui et al., 2017). The F1 coefficient, which combines precision and recall values, is recommended for overall performance, with a value of >0.70 considered adequate (van Rijsbergen, 1979). The Perspective API for identity attack narrowly surpassed (0.74) the threshold, making it an appropriate dependent variable for anti-refugee attitudes.

*Table 4: Pearson correlation matrix for Penally Perspective API scores*

	Toxicity	SevToxicity	Profanity	ID Attack	Threat	Insult
Toxicity	1	0.952	0.877	0.77	0.547	0.963
SevereToxicity	0.952	1	0.916	0.776	0.644	0.874
Profanity	0.877	0.916	1	0.605	0.475	0.796
ID Attack	0.77	0.776	0.605	1	0.478	0.731
Threat	0.547	0.644	0.475	0.478	1	0.419
Insult	0.963	0.874	0.796	0.731	0.419	1

Second, a Pearson correlation test was conducted<sup>15</sup> to determine the interrelatedness among the attributes. As shown in Table 4, most of the attributes exhibited moderate to strong correlations (above 0.6) with identity attack, except for threat. Gruzd and Mai (2021) recommend that when correlation scores among perspective API attributes are strong, researchers should focus on analysing only one attribute, especially if it is most relevant to their research question. Therefore, identity attack was chosen as the dependent variable to measure anti-refugee tweets related to the Penally asylum accommodation. Because of its low correlation scores with all other attributes, threat was included in the regression models as an independent variable.

---

<sup>15</sup> All correlation values are significant at the 0.01 level (2-tailed)

### **3.4.6. Preprocessing**

After collecting four data sources, the preprocessing phase was initiated and involved two stages: (1) data cleaning and (2) manual preprocessing. In the first stage, the fragmented data sources were combined, and daily counts from news headlines and Google searches were added to each tweet based on date timestamps. For the FOI police recorded data, month categories were created from timestamps. September, with the highest count of reported hate incidents, was chosen as the reference category for regression models. The second stage focused on cleaning data and correctly identifying variables such as hashtags. Due to the more accessible and research-friendly dataset format of Academic Twitter compared to NCapture, this stage was faster than the next chapter.

In the second stage of preprocessing, social factor predictors such as hashtag and agent types were established. Agent types were informed by the work of Williams and Burnap (2016) and Ozalp et al. (2020). Agent types including political, news, police, and activists, and lists were used to identify police and political agents while pattern matching was used for media agents. Pro and anti-refugee activists were identified by manual inspection of Twitter bio metadata, with inter-coder reliability (ICR) tests ( $\alpha = 0.98$ ) between the primary researcher and secondary coder, suggesting strong validity. Hashtag types were manually coded and included those that appeared five or more times in the dataset; as such, these referred to highly propagated hashtags in the dataset, reflecting wider social sentiment and avoiding multicollinearity with the hashtag number predictor. Pro-refugee hashtags with high prevalence included #LovePenallyHateRacism, #RefugeesWelcome, and #MigrantInvasion, while anti-refugee hashtags included #RefugeesNotWelcome. Tweets without a coded prevalent hashtag were coded as 'non-hashtag' and served as the reference category in the regression models. The ICR tests showed good levels of coder agreement ( $\alpha = 0.92$ ) for hashtag types, but some notable disagreements were observed, such as with the hashtag #CloseTheBarracks, which was initially interpreted as anti-refugee but was actually being used by pro-refugee actors who deemed the barracks to have unacceptable living conditions.

### **3.4.7. Methods of Analysis**

#### *3.4.7.1. Data Visualisation*

The analysis involves two phases: data visualisation and regression models. It is a predominantly quantitative study that incorporates qualitative components in the form of data visualisation. This is an embedded design approach, as Zhang and Creswell (2013) recommended. Big social data is recognised as a useful tool for extensive temporal research; however, it is often “noisy” and “unstructured”, which can impact interpretation (Williams and Burnap, 2016: 214). Therefore, data visualisation processes are employed to enhance interpretation and identify points of interest that may not be immediately apparent in raw datasets (Burnap et al., 2015: 82).

Visualisations allow for data aggregation, highlighting key themes while preserving the anonymity of individual social media users (Williams et al., 2013). Anonymity was particularly important, as individuals rather than organisations published most content<sup>16</sup>. Subsequently, the visualisation processes effectively identify key factors in the dataset, which can be further explored using regression models. During this phase, four forms of data visualisation are produced.

First, aggregated time-series line graphs are used to illustrate overall tweet frequency and anti-refugee content<sup>17</sup>. This approach helps to identify significant moments in the trigger event case, study information flow trends about the production of anti-refugee content, and confirm the Google search term determined issue attention cycle period, as previously mentioned. The second phase of analysis involved the use of data visualisation tools for a more in-depth exploration of the dataset. The online research tool CommuAnalytic was used to generate network visualisations demonstrating the dataset's reply networks (Gruzd & Mai, 2021). By filtering the reply network by anti-refugee content, initial interpretations of the role that online reply networks played in the spread and diffusion of anti-refugee content were made before further inspections in the regression models.

---

<sup>16</sup> For more information see ethical considerations

<sup>17</sup> Above the 0.7 identity attack threshold

The third visualisation in this study incorporated Google search terms as a supplementary dataset. Researchers encountered a methodological hurdle related to acquiring precise geo-tagged data from online discussions. This challenge arises because only 1% of users permit geo-tagging (Sloan et al., 2013). While some researchers have attempted to locate users through tweet and bio metadata (Ozalp et al., 2020), this study opted to use Google search data to identify online discussions related to the search term: "Penally Camps." Leveraging Google search data offered a larger sample size of online searchers compared to Twitter, thereby providing a richer source of data for analysis. An aggregated word cloud was used for qualitative insights into the main topics discussed, both in the exploratory phase and the discussion section after the regression models, to identify themes in anti-refugee tweets and between different agents.

### ***3.4.8. Regression Models***

#### *3.4.8.1. Independent Variables*

The study used the same independent variables in both tweet engagement and anti-refugee content models, grouped into four-factor categories. Content factors included five scale variables (mentions, hashtags, word count, threat, and sentiment) and the use of URLs, with identity attack also included as an independent variable for the engagement models using the Perspective API. Social factors were split into two categories: prevalent hashtags (pro- and anti-refugee hashtags) and social actors, including activists, media, police, and political agents. External factors included news headlines and Google searches. Control factors consisted of the eight months during which data collection occurred, all dummy coded. September was used as the reference category due to the highest recorded levels of hate crimes and incidents.

### 3.4.8.2. *Dependent Variables*

#### ***Engagement models***

Three distinct models were developed in this study, each utilising a different dependent variable to evaluate tweet engagement to measure online engagement. Due to the exclusion of retweets during the initial data collection phase, only original tweets were included, which prevented the measurement of tweet survival<sup>18</sup>, unlike in other studies employing a similar framework. While this represents a limitation, it also provided an opportunity to include other metrics for engagement, such as likes and replies, which have not been extensively studied in previous research. Data from three models were examined to assess differences in size engagement. Tweet engagement-dependent variables (retweets, likes, and replies) were skewed and over-dispersed, meaning a negative binomial model was fitted.

#### ***Anti-refugee content model***

The perspective API classifiers are unique because they employ scales rather than ordinal data for antagonistic speech (Ozalp et al., 2020; Gallacher, 2021). As a result, an OLS regression model was selected as the best fit for the data, and correlation analyses were conducted along with tolerance statistics to ensure there were no multicollinearity issues among the model's predictor variables.

## **3.5. Empirical Chapter 3**

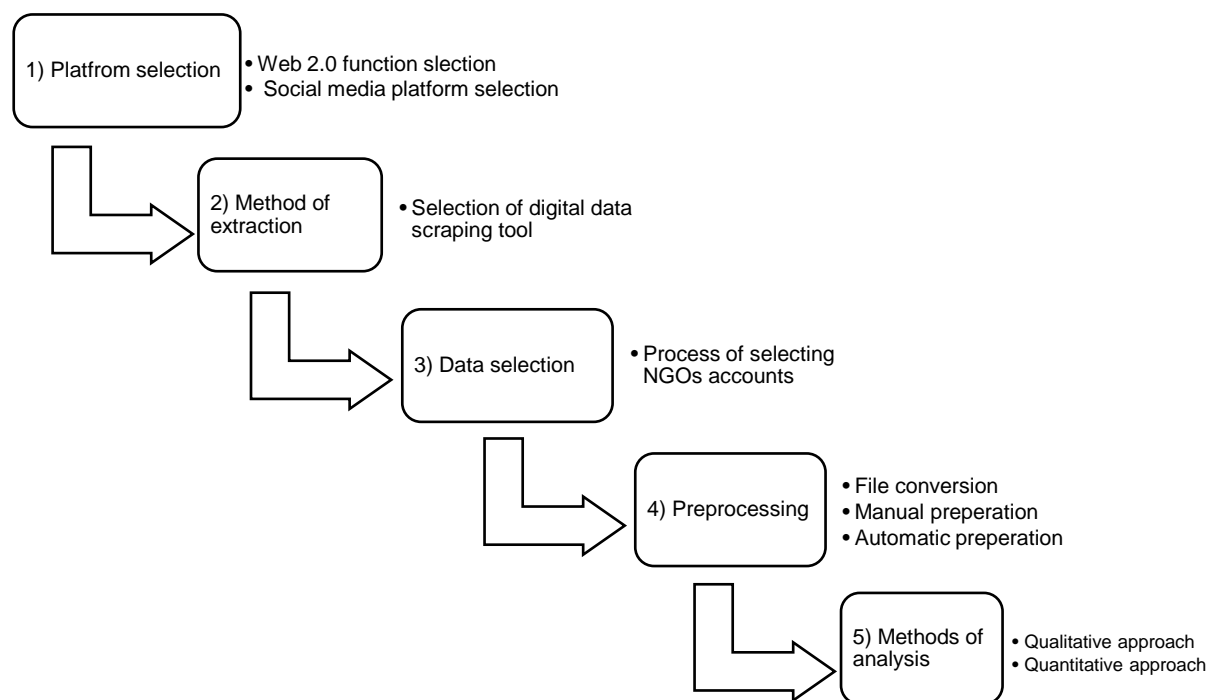
While social media use by public sector organisations has been extensively studied (Denef et al., 2013; Ferguson & Soave, 2020; Lieberman et al., 2013; Meijer & Thaaens, 2013; O'Connor, 2015; Panagiotopoulos et al., 2014; Cullen, 2022), there has been relatively less attention paid to the use of social media by NGOs (Guo and Saxton, 2010). This section furthers understandings of how non-profits use social media to combat antagonistic speech online and foster online cohesion and

---

<sup>18</sup> Tweet survival refers to the duration between a tweet's first and last retweet within a given time.

collective efficacy (Sampson et al., 1997), given their similarities in goals and objectives with public sector organisations, such as the police, particularly regarding community tensions.

In Chapter Six, communication strategies of minority rights NGOs in Wales is examined using content analysis. Stemler's (2000) definition of content analysis as a systematic and replicable technique for categorising text into fewer content categories is used. The study evaluates social media use by these organisations in promoting community cohesion and provides a framework for future studies. The study is based on Mergel's (2010) digital communication typology, informing the day-to-day practices of NGOs in promoting cohesion and collective efficacy (Sampson et al., 1997). Figure 8 depicts the five key steps of the research methodology, which are described in detail in this section.



*Figure 8: Methodological steps taken in Chapter Five*

### **3.5.1. Platform Selection**

Given the documented objectives of NGOs, such as expanding networks and increasing information reach (Van De Velde et al., 2015), non-reciprocal structures were deemed most appropriate for this study. Non-predetermined communications are essential to understanding networking strategies (Mergel, 2013) on social media. In addition, public space platforms offer research-friendly APIs for data collection, unlike many semi-public platforms (Dixon, 2022). For instance, Twitter provides API tools for data collection, enabling easier access to data at a meaningful scale (Burnap et al., 2015).

The selection process identified three public space, non-reciprocal platforms with significant user prevalence in the UK (Dixon, 2022): Twitter, Instagram, and YouTube. In choosing between them, modalities were considered. Modalities refer to the forms in which content is presented, such as images, videos, audio, and text (Waterloo et al., 2018). Research indicates that some social media platforms promote a balanced use of modalities, while others favour specific modalities (Kietzmann et al. 2011; Gligorić et al., 2018).

While Instagram and YouTube favour visual and audio-visual modalities, Twitter supports a more diverse range of content strategies (Waterloo et al., 2018). Since this study aims to examine a broad range of strategies and tools used by NGOs, Twitter was chosen. The selection of Twitter was based on the overarching pragmatist approach adopted throughout this thesis, which considered factors such as ease of data extraction, non-reciprocal networking structures, and the balanced provision of modalities. Therefore, Twitter was deemed the most practical option for data analysis<sup>19</sup>.

---

<sup>19</sup> For a more compressive overview of why Twitter was selected, see Appendix 4.

### **3.5.2. Method of Extraction**

An evaluation of the four extraction methods against the established criteria indicated that only N-Capture and Search API are suitable. As previously noted, these two methods are comparable in functionality, with the primary distinction being the accessibility of Search API, which necessitates a rigorous application procedure. In contrast, N-Capture is freely available and can be easily installed as a web browser extension. Given the modest sample size of tweets required for collection (n=5000), the additional capacity offered by Search API was deemed unnecessary. Consequently, the NCapture plugin for NVivo was selected for data extraction purposes.

### **3.5.3. Data Selection**

With the social media platform and data extraction method determined, the next step was to identify the relevant third sector accounts. As there is no pre-existing list of third sector minority rights organisations in Wales, organisations were chosen based on their affiliations with the Hate Crime Criminal Justice Board and Charities Against Hate (2020). After compiling a comprehensive list of NGOs, a selection criterion was established: (1) the Twitter account must be active, with frequent tweets posted on a daily or weekly basis, (2) the Twitter accounts must exclusively operate within Wales, (3) the Twitter accounts must exist to represent marginalised communities relating to protected characteristics, and (4) the Twitter accounts must belong to NGO, rather than individuals.

Finally, ten accounts were randomly selected, with an equal representation of focus<sup>20</sup> and organisational size. The final selection consisted of ten separate Twitter accounts belonging to NGOs (summarised in Table 5). The collection encompassed both tweets and replies made by the accounts. Replies were crucial for evaluating the organisations' networking strategies through interaction with other accounts (Mergel, 2012), while retweets were excluded due to their problematic nature for

---

<sup>20</sup> Focus relates to each of the nine protected characteristics (Home Office, 2010)



content analysis (O'Connor, 2015; Cullen, 2020). Retweets by external stakeholders may compromise data validity, as this study focuses on communication strategies employed by NGOs.

*Table 5. Twitter account summary*

Twitter Account <sup>a</sup> (Case)	Code	Followers <sup>b</sup>	Following <sup>b</sup>	Size (code)	Tweet Frequency
EYST n=500	1	5,040	3,886	Medium (2)	High
Pride n=500	2	16,402	5,984	X-Large (4)	Medium
Race Equality First n=500	3	4,812	4,534	Medium (2)	High
Stonewall n=500	4	9,834	1,531	Large (3)	Low
Trans Aid Cymru n=500	5	878	123	Small (1)	High
WomenConnect First n=500	6	2,799	873	Small (1)	Low
Welsh Refugee Council n=500	7	4,434	835	Medium (2)	Low
WelshWomensAid n=500	8	11,109	2,012	X-Large (4)	Medium
Diverse Cymru n=500	9	9,439	2,574	Large (3)	Low
Disability Wales n=500	10	14,094	2,068	X-Large (4)	High

*Tweets included cover 10th November 2018-10th February 2021. Total n=5000. As of download date (10<sup>th</sup> February 2021).*

Data for this study was collected from Nov 10th, 2018, to Feb 10th, 2021, with a sample of 5,000 tweets. Previous research on public-sector digital communication used smaller samples (Denef et al., 2013; Lieberman; O'Connor, 2015), whereas some research on third sector digital communication had larger samples but focused on content factors rather than specific strategies (Guo & Saxton, 2018; Haplin et al., 2021). Among limited research on third sector digital communication strategies, most studies used samples of less than 1,000 tweets (Deschamps and McNutt, 2014; Guo and Saxton, 2014). Given the limited literature on third sector online communication strategies and the need for manual coding, this study selected a practical sample size of 5,000 tweets instead of larger sample sizes (e.g., 4,340,891 in Haplin et al., 2021), which could be impractical for a PhD candidate with limited resources.

### 3.5.4. Preprocessing

During the preprocessing phase, an effort was made to convert certain qualitative data into numerical format for analytical purposes. The preparation of data was carried out in three distinct steps: file conversion, manual preparation, and automatic preparation. The categorisation of hashtags and emojis was achieved through a combination of both manual and automatic procedures.

#### 3.5.4.1. Manual preparation

In the first preprocessing phase, manual coding was conducted to categorise tweets into both broad and sub-strategies, along with coding other tweet-related factors. This manual processing involved two distinct phases. The initial phase adhered to Mergel's (2010) communication strategy framework and created a "strategy" variable to classify tweets into push, pull, networking, or transactional strategies. Dummy variables were generated for each strategy category. Categorising each tweet into a specific strategy category was essential for conducting content analysis and evaluating engagement levels, including metrics such as likes, retweets, and replies (Boyd et al., 2010). Additionally, it enabled the assessment of discrepancies in content modalities employed between the strategies. Please refer to Figure 9 for a summarised overview of the coded strategies in this study.

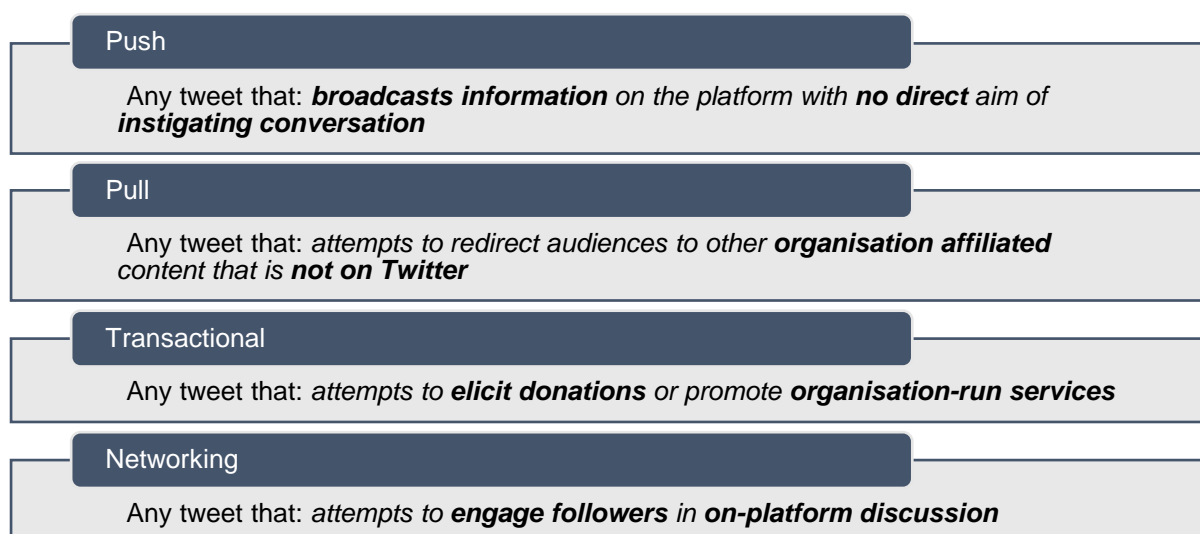
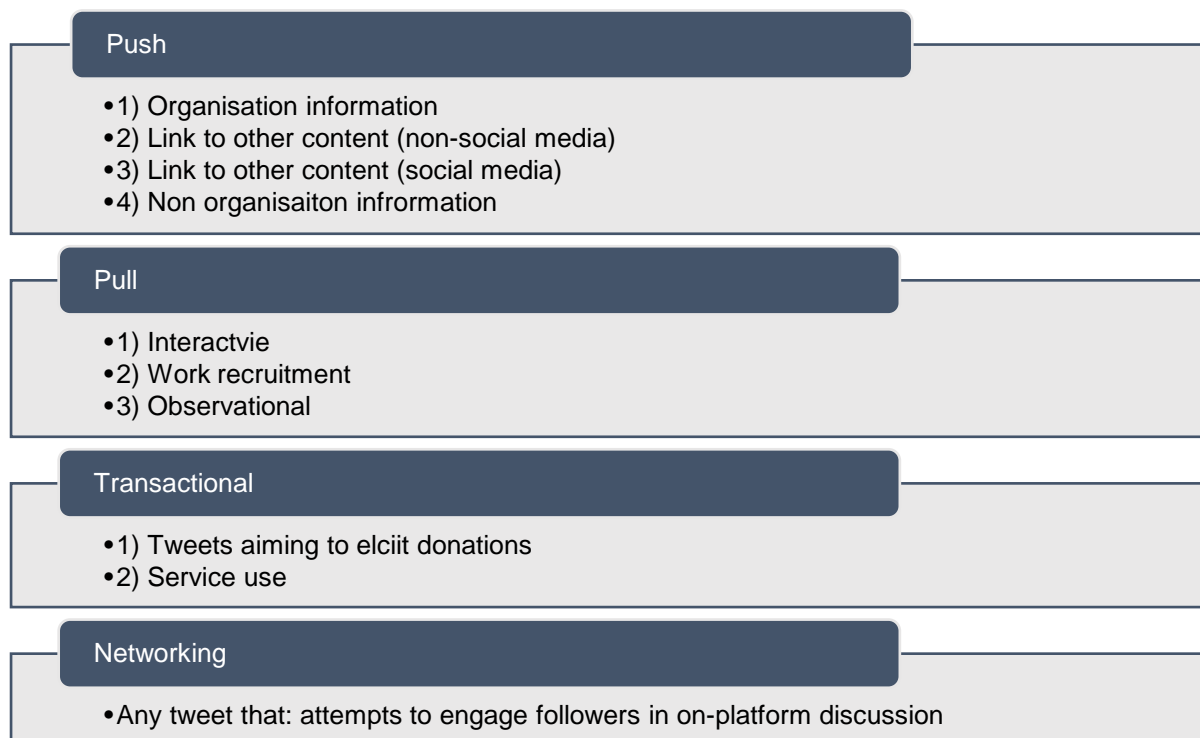


Figure 9: Coding indicators for broad strategies

Each tweet was categorised using Mergel's framework, and three dummy variables were manually coded to capture other focus, intervention, and celebration. The "other focus" dummy variable indicates tweets that support a protected characteristic other than the organisation's primary focus.

The "intervention" dummy variable identifies informal social controls for discriminatory attitudes in digital contexts as defined by Sampson et al. (1997) and Groff (2015). Therefore, Intervention tweets publicly condemned specific attitudes, whether online or offline. They must pertain to one particular case rather than general discussions of discrimination or inequality. The "celebration" dummy variable includes days, weeks, or months of celebration, mostly related to minority groups, but also includes national holidays such as St David's Day for consistency during coding. The broad communication strategies were tested for inter-coding reliability (ICR) before proceeding to the second phase of manual coding. Details on this test will be provided later in the section.

Communication strategies were categorised using Mergel's framework, but the coding phase also used an open approach (Corbin & Strauss, 2015) to gain a deeper understanding of the implementation of each strategy. Findings were consistent with previous studies (Meijer and Thaens, 2013; O'Connor, 2015), revealing that sub-strategies could be distinguished within Mergel's broad push, pull, and transactional strategies. Networking communication lacked distinct sub-strategies due to its decentralised and pluralistic nature (Meijer and Thaens, 2013; Mergel, 2013). Nine sub-strategies were inductively categorised within the same broad strategies, including four for push, three for pull, and two for transactional (see Figure 10).



*Figure 10. Digital communication sub-strategies*

The second manual coding phase excluded tweets with the networking strategy (n=650) because no discernible sub-strategies were identified. Using the notes made during the first coding phase, the coding manual was expanded, outlining the criteria for each sub-strategy (see Appendix 6). The sub-strategies for the push strategy include organisation information, non-organisation information, link to other content (non-social media), and link to other content (social media).

Organisation information sub-strategy refers to tweets mentioning the organisation's work or operations, while non-organisation information sub-strategy pertains to tweets discussing concepts unrelated to the organisation's operations. The link to other content (social media) sub-strategy incorporates a more dialogical approach, mentioning other nodes while broadcasting information. It still operates as a push strategy but acknowledges other entities while disseminating information. The link to other content (non-social media) sub-strategy disseminates information by providing links to non-organisation webpages, such as news articles. The sub-strategies for the pull strategy are interactive, work recruitment, and observational. The interactive sub-strategy encourages audience participation in organisation activities via tweets.

The work recruitment sub-strategy aims to recruit new staff or volunteers through tweets. The observational sub-strategy involves prompting followers to visit the organisation's website by following a hyperlink in the tweet, rather than engaging in activities. Finally, the transactional sub-strategies included efforts to elicit donations or promote organisation-run services.

### **3.5.5. ICR Coding Procedure**

Content analyses identify common themes in textual data (Stemler, 2000). ICR tests assess the validity of manual coding procedures in content analysis (Lombard et al., 2002). Krippendorff's alpha coefficient, ranging from 0 to 1, tests multiple coding categories (Gwet, 2011). A statistic above 0.67 indicates sufficient agreement, while above 0.8 is considered high intercoder agreement (Antoine et al., 2014; Carletta, 1996). To validate manual coding (Bryman, 2012), a second researcher performed inter-coding on 2,500 tweets, evaluated using Krippendorff's alpha statistic, demonstrating high intercoder reliability<sup>21</sup>. The study found high consistency in manual coding between the two coders, as shown by Krippendorff's alpha values (Lombard et al., 2002), with a score of 1.00 for the other focus variable indicating complete agreement. Despite minor disagreements, Krippendorff's alpha scores remained above acceptable thresholds, demonstrating coding reliability. Common coding disagreements involved Mergel's (2010) communication strategies, particularly confusion between "push" and "pull," due to misunderstandings about the nature of URLs as external pages versus affiliated pages.

Additionally, disagreements in the celebration category were due to secondary coders not identifying references to celebrations in hashtags (e.g., #IWD2020). Differences in perspective regarding the relationship of interventions to specific events or tensions in a general sense were the main cause of disagreements in the intervention category. Sub-strategy coding revealed disagreements between link to other content (non-social media) and observational pull, as well as between link to

---

<sup>21</sup> The Krippendorff's alpha coefficient for each item was as follows: broad communication strategies ( $\alpha= 0.95$ ), celebration ( $\alpha= 0.91$ ), intervention ( $\alpha= 0.81$ ), other focus ( $\alpha= 1.00$ ), and sub-strategies ( $\alpha= 0.83$ )

other content (social media) and networking due to coders misinterpreting links and mentions of other Twitter accounts. Nevertheless, high Krippendorff's alpha scores (0.81 to 1.00) demonstrate the inter-coding process was reliable, suggesting that human communication's subjective nature influenced the differences observed (O'Connor & Joffe, 2020).

### **3.5.6. Automatic Preprocessing**

The automatic phase involved data cleaning and generating new variables using available information. The disorganised data required thorough cleaning to make it accessible for analysis. Automatic processes produced variables not provided by NCapture. First, the account factors (tweet frequency and account size) were devised automatically using SPSS. This was achieved by utilising the account variable for each of the ten accounts. Accounts with 0-3,000 followers were classified as 'small' (1), 3,000-6,000 followers as 'medium' (2), 6,000-10,000 followers as 'large' (3), and 10,000+ as 'x-large' (4). The same process was used to classify tweet frequency, with accounts that had a tweet per day ratio of 0-1.5 classified as 'low' (1), 1.5-2.5 as 'mid' (2), and 2.5+ as 'high' (3).

After establishing account factors, the SPSS file was converted to excel for further automatic preprocessing. Excel functions were used to establish variables needed to address research questions. Hashtags, word count and mentions were converted from string variables to scales<sup>22</sup>. The 'URL' variable was generated<sup>23</sup> and recoded into a numeric dummy variable with '1' representing 'TRUE' (hyperlink present) and '0' for 'FALSE' (hyperlink absent). The same process was applied for the mention covid variable to the following terms: 'covid', 'corona', 'covid19', 'lockdown', 'corona virus', 'c19' and 'vaccine'. The resulting variables were then merged into a dummy

---

<sup>22</sup> These were converted to numeric data using a word count formula employing 'LEN', 'TRIM', and 'SUBSTITUTE' commands (Parkes et al., 2013).

<sup>23</sup> The =ISNUMBER function in excel (Holding & Martin, 1998) was used to identify hyperlinks (URL) in tweets, generating outcomes of 'TRUE' or 'FALSE'

variable<sup>24</sup>. This provided an objective measure, meaning that ICR tests were not necessary. Time of the day was automatically coded, with the morning representing 6:00 am – 11:00 am, daytime 11:00 am – 5:00 pm, night 5:00 pm – 10:00 pm and evening 10:00 pm – 6:00 am. A dummy 'weekend' variable was created with weekdays (0) and weekends (1).

### ***3.5.7. Hashtags and Emojis***

The final coding stage combined manual and automated techniques to classify hashtags and emojis. Emojis and hashtags were coded once due to their repetition. Emoji categories were established inductively, except for group-specific emojis guided by the "Diversity Language Model" (Swartz et al., 2020). All emojis were listed using a character frequency tool and filtered by letters and symbols. Four main uses of emojis were identified: (1) emotive, (2) group-specific, (3) activity/weather, and (4) symbolic. ICR procedures achieved complete coder agreement, likely due to the small number of emojis in the dataset (n=50) and their easy interpretability. Correlation tests showed no overlap between emoji types, allowing for inputting of emoji types in regression models about a "no emoji" reference category. Table 6 presents detailed definitions, frequent examples, and usage in context for each of the four emoji types.

---

<sup>24</sup> The dummy coding was '1' if it included one or more terms and '0' if none were mentioned.

Table 6: *Emoji typology*

Emoji Type	Definition	Frequent Study Example	Example of Usage
Emotive	An emoji used to emphasise and/ or express an emotion of the author		"@beefanddairy Thank you so much
Group Specific	An emoji that references a particular social group or movement. These were selected based on the "Diversity Language Model" (Swartz et al., 2020: 95)- see chapter XXX		" #TransRightsAreHumanRights <a href="https://t.co/7TnGlbLUUq">https://t.co/7TnGlbLUUq</a> "
Activity/ Weather	An emoji that promotes an activity, sometimes using weather, discussed in any given tweet. These are often used when promoting an organisation-based activity.		"Have you booked on to our FREE training yet? For more information and to book you place, visit our website <a href="https://t.co/wV8mu28zuZ">https://t.co/wV8mu28zuZ</a> "
Symbolic	An emoji used to symbolise an action. For example, phone emojis were used when promoting a call service, or the down arrow key for downloading an online resource		"Did you know? Our #BAMEHelpline  is open as normal Mon-Fri 10:30am-2:30pm. Support available in different languages!"

The "=ISNUMBER" function generated dummy variables for the identified emoji types. A reference category of "no emoji type" was included for cases without an emoji, and cases with multiple instances of the same emoji were coded as "1." The hashtag types were coded using a framework from previous studies (Saxton et al., 2015: 161), with some exceptions discussed later in the section.

The word frequency analysis revealed the existence of 1464 unique hashtags in the study, and hashtags used four or more times were coded for analysis. The ICR test showed a reasonable level of agreement (0.78), with most disagreements arising between organisation and external branding, where coders disagreed on the affiliation of the hashtag. A correlation analysis revealed many instances of overlapping hashtag types in single tweets. Rather than determining the most salient hashtag type in each tweet, cases with two or more hashtag types were all coded. The reference category of "no hashtag" was therefore not included in the models. Once again, the "=ISNUMBER" function produced dummy variables for each hashtag type. Table 7 presents the definitions established by Saxton et al. (2015: 161) and examples from this study's dataset.

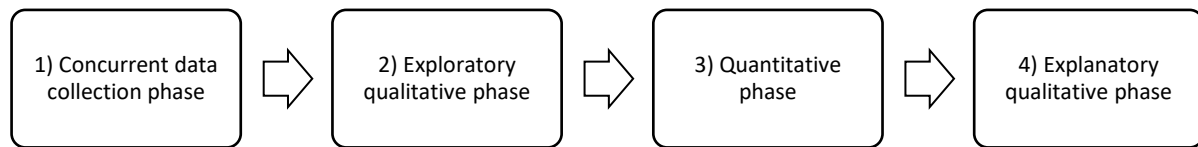


Table 7: Hashtag typology based on Saxton et al. (2015: 161)

Hashtag Type	Definition	Frequent Example
Knowledge/ Informative	Any hashtag that seeks to help inform audiences on concepts and terms associated with the organisation's work	#LGBT #covid19 #hatecrime #dabill #euss #brexit #bame #womensrights
Events	Events that are either organisation run, or a local, national or global day/ week of recognition	#blackhistorymonth #hatecrimeawarenessweek #holocaustmemorialday #worldmentalhealthday #blackouttuesday
Call-To-Action	"hashtags can be used to mobilize audiences for collective action, whether to engage in direct online or offline action or simply to assist in further disseminating its public education messages"	#read #comeoutvoting #transcrowdfund #togetherwecan #standwithsurvivors #reachout
Values and Goals	Organisational values or goals. Useful for reinforcing the organization's core values and ultimate strategic goals	#refugeeswelcome #womensupportingwomen #stophate #eurewelcome #healthyandactive #nationofsanctuary
Dialogic	'Chat' hashtags, prompting audience to interact in discussion	#askmecardiff #boreda #haveyoursay #askme #thankyou #timetotalk
Time and Place	Hashtag relating to time or location	#wales #cardiff # butetown #swansea #cathays #newport #roath
Business	Related to business issues, stocks, companies, etc. Captures a wide range of non-health, non-advocacy-related hashtags, used in the sample	No hashtags were considered business related, any mirroring Saxton et al (2015) were coded into the branding hashtags
Organisation Branding	Organisation-specific hashtags, unique organization identifiers, hashtags noting one of the organization's program names, this includes support services	#WomenConnectFirstStaff. #BAMEHelpline #WCF #PrideCymru #WRCPlaySuppor
External Branding	The same criteria as organisation branding, but for an organisation or movement autonomous of the author organisation	#BelongingProject #YourPride #worldcafeonwheels #blacklivesmatter #RainbowLacesDay #DWMManifesto

### 3.5.8. Methods of Analysis

The analysis adopted a mixed-method approach following the pragmatist epistemological perspective of the thesis. While a concurrent approach was utilised during the data collection stage, as is standard with social media data (Snelson, 2016), the analysis employed a multiphase sequential approach, (Creswell & Plano Clarks, 2011), with three distinct analytical phases:



*Figure 11. Analytical phases for Chapter Six*

The first phase of the analysis utilised methods commonly employed in previous research, with Mergel's (2010) framework as the foundation for identifying specific sub-strategies via exploratory content analysis (Meijer and Thaens, 2013; O'Connor, 2015). Four broad categories, reflecting Mergel's typology, were predetermined, and an additional nine sub-strategies were inductively derived during the qualitative exploratory phase. This phase involved a meticulous examination of tweet extracts from the dataset to discern the presentation and functions employed in different strategies. Examples were provided to illustrate key differences in communication strategies and how they vary. This approach facilitated a deeper understanding of NGOs' strategies to promote inclusive content and diffuse on social media. This study diverges from similar works, such as O'Connor (2015), by focusing on the general usage of approaches rather than differences between specific organisations. While qualitative insights into communication strategies offer important insights, some academics argue that a singular approach could be limiting in obtaining comprehensive understandings. They advocate for incorporating quantitative analysis, which provides a more holistic, exploratory, and explanatory examination of the various strategies employed in content analyses (Williams & Shephard, 2017; Cullen, 2022).

Previous studies have used quantitative methods, such as regression models, to link organisational factors to strategy use (Cullen, 2022). While acknowledging the importance of a mixed methods approach, this study deviates from previous quantitative methods by examining two dynamics: (i) differences in content factor prevalence amongst strategies using Kruskal Wallis tests, and (ii) which communication strategies predict audience engagement measured with regression models. This study is the first to examine audience engagement in relation to Mergel's (2010) online communication strategy typology. However, the findings are limited to NGOs and are not generalisable to other organisation types. Regression

models were selected with careful consideration of the implications of using count data. Once again, negative binomial models were applied to the over-dispersed engagement data (Wulu et al., 2002).

This study examined additional forms of engagement, such as likes and replies, to identify any significant differences in their outcomes. Independent variables were categorised into five-factor groups: content, strategy, account, emoji/hashtag types, and controls. Tweet survival was deemed less applicable to this study due to non-event-specific content (Williams and Burnap, 2016; Ozalp et al., 2020). The three models (likes, retweets, and replies) explain different forms of audience engagement, and predictor coefficients were cross-examined to identify their key differences. Previous studies on audience engagement have primarily focused on retweets as the sole dependent variable (Boyd et al., 2010), but this study adds to the literature by examining additional forms of engagement.

The quantitative phase of the study identified discrepancies between strategies in terms of content factors and audience engagement. This phase contributes to the identification of how sub-strategies contribute to core aims, as indicated in the literature review. The outcomes of the quantitative phase were used to categorise sub-strategies and identify how they contribute to the five broad aims of promoting inclusive content and achieving collective efficacy. These broad aims include (i) forging partnerships and establishing broad networks with other organisations, (ii) actively combating antagonistic speech and promoting inclusiveness, (iii) disseminating positive, diverse, and inclusive content, (iv) enhancing their digital transactional environment, and (v) continuously gathering information to inform their initiatives. The final qualitative phase of the study provides a deeper understanding of the relationships between the sub-strategies and the fundamental aims, considering the context in which content factors are used and why they lead to higher engagement levels.

### **3.6. Ethical Considerations**

Ethical practices are increasingly emphasised across various disciplines (Farrimond, 2012), particularly in the social sciences, which investigate human behaviour and social systems. Ethical considerations in research involving human subjects are crucial for safeguarding participants' rights and well-being (Birnbacher, 1999). In this thesis, strict ethical standards were adhered to minimise harm to all involved, including the researcher, participants, and others. This included obtaining ethical approval from the Cardiff School of Social Sciences Research Ethics Committee (see Appendix 7), following the guidelines of the Economic and Social Research Council framework (ESRC, 2019), and completing Cardiff University's module-based training procedure before conducting any research.

Feedback from the ethics committee prompted significant enhancements to the research design. This was most notable in the consent and information sheets used in Chapter Four. Participants received information sheets outlining the survey content, data handling procedures, and study purpose (see Appendix 2). Consent forms were also used to reinforce the provided information and allow participants to opt-out at any time (see Appendix 3); for the second empirical chapter, which employed social media data from organisations, information sheets and consent forms were unnecessary. Differences in social media user consent are explored later in this section.

To address concerns raised during the review about consent confirmation in digital surveys, the researcher implemented a dual approach involving both traditional physical signatures on scanned forms and electronic signatures. This decision was reached following consultations with the review panel and my supervisor, who expressed concerns that relying solely on survey completion as an indicator of consent might not sufficiently mitigate anxieties about robust consent processes. Participants were offered the flexibility to either electronically complete surveys and submit scanned signed forms or utilise Qualtrics' e-signature feature directly. This approach not only ensures compliance with digital data collection standards but also helps reduce the likelihood of non-responses due to potential deterrents such as the inconvenience of scanning signatures. After the ethics review, adjustments were also

implemented to enhance clarity regarding the opt-out process. This involved providing explicit instructions and information about how participants could withdraw from the study if they chose to do so. This clarity ensures that participants understand their right to withdraw at any time without repercussions, thereby promoting transparency and respecting their autonomy in the research process. To assist in doing so, contact details for not only the researcher but also the supervisor were included to facilitate continuous communication and address participant concerns promptly. These revisions aimed to improve transparency and participant engagement throughout the study.

Despite the differences among the empirical chapters, ethical considerations were consistently maintained, including the need to maintain confidentiality following the General Data Protection Regulation. Methods to maintain confidentiality varied, with the first empirical chapter involving a survey analysing the cohesion delivery network in Wales and the extent of cooperation among network nodes. Maintaining confidentiality in research can pose challenges (Wiles et al., 2008). The first empirical chapter analysed the cohesion delivery network in Wales, where preserving anonymity was crucial due to the central role of participant identity in the network analysis. In Chapter Four, adjustments were likewise made to the survey methodology following the review advice to remove regional analysis. This modification aimed to safeguard participant anonymity after it was advised that regional data could potentially identify specific respondents due to the low coverage of certain roles in some areas. The researcher balanced the need for data with anonymity by comparing general themes among cluster groups instead of attributing information to individual participants. For smaller groups like academics and cohesion coordinators, where only one coordinator and two officers existed per region, data was not presented by region. To ensure anonymity, individual statements were identified with cluster group codes instead of participant names (Braun & Clarke, 2013). Extracts were redacted if they could reveal identifying information, such as location.

The ethical considerations for the first empirical chapter were easier to address due to the availability of a rich body of literature on ethical issues related to traditional (Oldendick, 2012) and online survey methods (Evans and Mathur, 2018). However,

research using big data faces a significant ethical dilemma, as existing frameworks tend to reflect the practices established in the 20th century and fail to consider emerging factors associated with digital society (Pastor-Escuredo & Vinuesa, 2020). This posed potential challenges for the analysis carried out in the second and third empirical chapters.

To address ethical challenges related to the handling and analysis of social media data, the researcher followed guidelines provided by the Cardiff University Social Data Science Lab, which is based on the Economic and Social Research Council framework (ESRC, 2019) (Figure 12). The six guidelines included compliance with the ESRC framework, obtaining ethical approval from the University, and implementing secure storage measures for datasets. These principles aligned with measures taken for non-social media data and formed the basis for ethical considerations in the second and third empirical chapters (Social Data Science Lab, 2021).

To adhere to ethical standards in social media research for publishing the content of posts, two approaches for obtaining consent, opt-in and opt-out, were used (Social Data Science Lab, 2021). Opt-in was implemented for sensitive tweets or vulnerable users, while opt-out was used for all other posts and users. Consent was obtained through on-platform instant messenger services and consent procedures, mainly for counter agents, as obtaining consent from anti-refugee propagators was challenging.

In addition, regular check-ins with my supervisor, who possesses expertise in social media data analysis extending beyond traditional ethical considerations, served as critical audits to ensure meticulous adherence to ethical guidelines and best practices in data handling and analysis. These periodic reviews not only bolstered methodological rigor but also ensured that data security measures were robust and aligned with evolving ethical standards throughout the research process.

My supervisor's deep understanding of social media dynamics and ethical implications was invaluable in navigating complex ethical landscapes associated with digital data. These discussions and audits enabled proactive adjustments to

methodologies, ensuring ethical integrity was maintained while maximising the depth of insights derived from social media sources.

For example, one of the significant challenges encountered was effectively balancing the extraction of insightful individual data from social media with stringent ethical constraints concerning participant privacy. To address this challenge, rigorous data aggregation techniques were implemented based on my supervisor's recommendation during a review when we discussed how to interpret and include specific data. These techniques were designed to identify and highlight broad thematic patterns and trends without compromising the anonymity or confidentiality of individual participants. This methodological approach not only safeguarded participant privacy rights but also facilitated the generation of meaningful insights from aggregated data. This was showcased in the qualitative themes among anti-refugee nodes in Chapter Five while maintaining anonymity when word clouds and anonymised network visualisations were employed (Williams et al., 2017: 1158). These formats provided insights into anti-refugee sentiments without compromising anonymity (Social Data Science Lab, 2021; Williams et al., 2017).

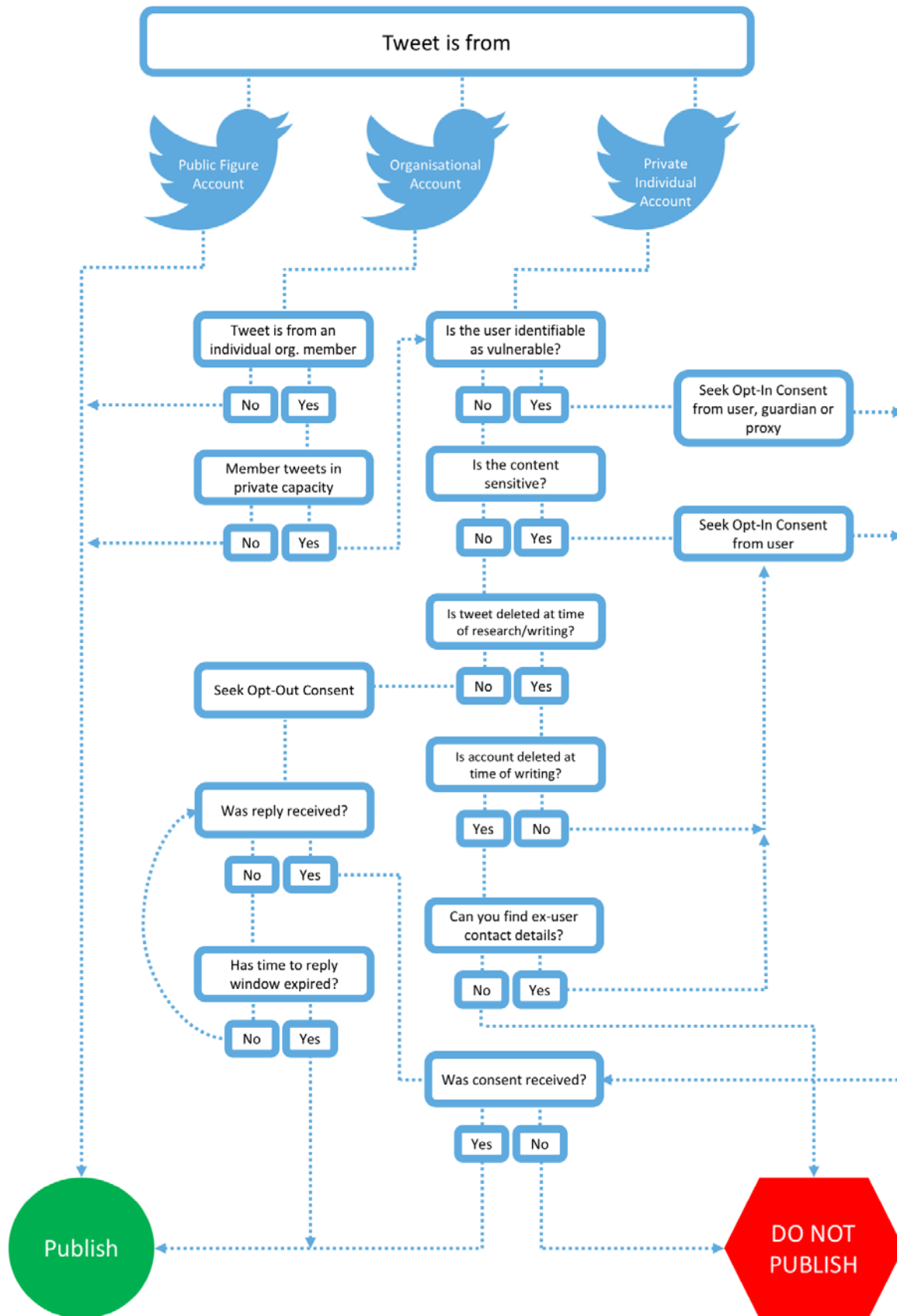


Figure 12: Decision flow chart for publication of Twitter communications  
(Williams et al., 2017: 1163)



The final principle follows Williams et al.'s (2017) flowchart for citing online accounts belonging to public organisations. This is relevant to the second empirical chapter, which analyses how NGOs use social media to promote inclusive content. According to the guidance, consent is not required for tweets from public organisations like government departments, law enforcement, local authorities, companies, and public figures like politicians (Social Data Science Lab, 2021). Thus, specific posts from these organisations were included frequently throughout the chapter in line with the lab guidance. This facilitated a greater emphasis on qualitative extracts, resulting in a more comprehensive examination of the communication strategies used.

### **3.7. Summary**

This chapter has provided a comprehensive overview of the research methods and datasets employed to investigate the strategies employed by stakeholders in Wales to achieve community cohesion. The research approach was firmly grounded in mixed methods pragmatic epistemology, emphasising the need to utilise various methods to comprehensively understand complex social phenomena.

It commenced with an overarching introduction to the mixed methods approach, establishing the methodological framework. Subsequent sections focused on specific study designs for three empirical chapters. The first chapter utilised primary survey data from the Welsh CDN to examine how stakeholders addressed tension-related issues. The second chapter analysed the communication strategies employed by third sector minority rights organisations using social media data. Chapter Five delves into the propagation of antagonistic attitudes and behaviours that stir up tensions through social media data, particularly after a case study-triggered event. Finally, this chapter delineated the ethical deliberations underpinning this thesis, elucidating the specific guidelines and policies adhered to. The next three chapters will delve into the original research carried out in this thesis that relates to the methodologies outlined.

## Chapter 4

### Mapping the Cohesion Delivery Network in Wales

---

#### 4.1. Introduction

Nodal governance has emerged as a potential solution for addressing social threats in contemporary society (Crawford, 2002). This multi-agency approach has gained recognition in recent academic discussions and has been identified as a prospective strategy for addressing community tensions (Chakraborti, 2018). This study builds on this momentum by investigating the emerging network of stakeholders involved in cohesion delivery in Wales - the cohesion delivery network (CDN). An emphasis is placed on cooperation frequency, quality patterns, and discrepancies in stakeholder priorities. The analysis draws on survey data collected from seven key stakeholder groups identified by the Welsh Government framework for action, including cohesion coordinators, cohesion officers, the Welsh government, community groups, academics, the police, and charities. Collective efficacy can serve as a conceptual framework to investigate community cohesion, as suggested by Sampson et al. (1997). When community members come together to tackle a shared issue, they develop a sense of collective efficacy. This shared belief in their ability to solve problems and effect positive change in their community can help strengthen social bonds and trust among community members. Sampson et al. (1997) propose that collective efficacy can be achieved through informal mechanisms and formally or externally induced actions.

Informal mechanisms exist at the community level and refer to mutual understanding, trust, and community willingness to intervene and support each other. As such, informal mechanisms are fundamental to achieving collective efficacy (Sampson et al., 1997). However, informal mechanisms are not mutually exclusive from external mechanisms, as community attitudes and behaviours can influence actions taken in external multi-agency partnerships and vice versa. Scholars have emphasised the importance of meaningful community integration in multi-agency

partnerships, termed "denizenship," to promote community cohesion and collective efficacy (Arnstein, 1969; Shearing and Wood, 2003; Kagan and Duggan, 2011). To provide an understanding of the network dynamics, this chapter examines cooperation frequency and quality among stakeholders and explores the underlying reasons for any patterns. This involves drawing on qualitative inputs from network stakeholders to provide explanatory accounts for why certain patterns exist. As such, this study offers a more nuanced and context-specific understanding of the dynamics of the CDN in Wales. In addition, this chapter examines the impact of the COVID-19 pandemic on day-to-day operations and cooperation patterns. The pandemic had significant implications for social and community life, including the delivery of cohesion initiatives. Therefore, this analysis explores the lasting changes the pandemic may have had on these stakeholders as a legacy. Finally, initial insights into social media's role in community tensions are explored, with stakeholders outlining the requirements for any potential tension monitoring tool that could be implemented.

## 4.2. Hypotheses

*H1. Evidence of a multi-agency (nodal governance) approach will be evident in the CDN, with the government taking a central role.*

This assumption is based on theorisations of a nodal governance approach to contemporary policing (Shearing, 2001). Although most surrounding literature suggests that broad agenda development remains at the state level, such as the Welsh government and similar institutions, most modern networks involve a wider range of stakeholders across sectors (Crawford, 2002). However, unlike many other modern networks, community-based issues tend to involve the private sector less, and instead, a greater emphasis is placed on community groups. Community engagement is, therefore, pivotal and, to occur, requires a willingness and capacity to participate from not only communities but other pre-established partners (Myhill, 2006).

*H2. Trends in the network will appear between stakeholder groups: (i) clustering of stakeholder groups will be evident (in relation to cooperation frequency), (ii) inter-cluster cooperation will be uncommon, (iii) non-reciprocal cases of cooperation will be found, and the extent will be greater in inter cases than intra.*

These assumptions are based on previous research that uses similar techniques, such as PCA and MDS (Levi and Williams, 2013). Levi and Williams (2013) set out to examine multi-agency partnerships in cybercrime reduction and included twelve stakeholder groups. Although the network of interest differs from the one explored in this study, a large amount of overlap exists in stakeholder groups. In relation to it is hypothesised that (i) clustering of stakeholder groups will be evident in MDS and PCA findings. (ii) relations between non-clustered stakeholder groups<sup>25</sup> (inter-meta) will be weaker and less frequent than intra-meta relations. This hypothesis is based on fundamental assumptions of PCA techniques suggesting that member groups of found components share commonalities and are more likely to correlate (Suhr, 2005). In relation to (iii), non-reciprocal cooperation cases will be found. This assumption is based on non-reciprocal relationships being a commonly discussed barrier to cooperation in multi-agency networks in surrounding literature (Cheminais, 2009).

*H3. Cooperation quality will predict cooperation frequency.*

This assumption is based on a study that examined predicting factors of cooperation frequency in a cyber-security multi-agency partnership. Node perceptions of other stakeholders had an effect on cooperation frequency. More specifically, if respondents deemed a particular agency as ineffective, they were more likely to cooperate with alternative clusters (Levi and Williams, 2013). Node perceptions of other stakeholders accounted for the most variance in cooperation frequency than any other identified sub-factor. Therefore, it is hypothesised that cooperation quality will have a positive association with cooperation frequency, and the extent of this will be greater within PCA-determined meta-clusters.

---

<sup>25</sup> As determined by PCA/MDS procedures

*H4. Node characteristics will predict cooperation frequency.*

This assumption is based on partnership literature that indicates that cohesive agendas are a key factor in improving cooperation (Kean and Hamilton, 2004). Specific stakeholder objectives (or characteristics) in this study include EU transition fund recipients, government-association, and agency prioritisation on detection, prevention, and response. Detection, prevention, and response agendas reflect priorities set out in the CDN agenda. It is therefore hypothesised that EU-transition-fund recipients will be positively associated with high cooperation frequency in cluster groups also receiving funding. Conversely, they will be negatively associated with non-recipient cluster groups. Likewise, government-based respondents' cooperation frequency will be highly associated with other government node groups, with non-government cluster groups being negatively associated. Finally, based on Cheminais (2009), it is hypothesised that agency agendas such as prevention, detection and response focus will also predict cooperation frequency.

### 4.3. Results

*Table 8: Descriptive statistics of CDN respondents (n=59)*

Independent Variables	Coding	n/ %
Cluster Group	Cohesion Coordinator	8 (13.6%)
	Cohesion Officer	8 (13.6%)
	The Welsh Government	6 (10.2%)
	Community Groups	10 (16.9%)
	Academics	7 (11.9%)
	The Police	9 (15.3%)
	Charities	11 (18.6%)
Regional	Two or More Regions	25 (42.4%)
	Single Region	34 (57.6%)
Government Based	Yes	16 (27.1%)
	No	40 (67.8%)
	Don't Know	3 (5.1%)
EU Transition Fund Funding	Yes	13 (22.0%)
	No	42 (71.2%)
	Don't Know	4 (6.8%)
Length of Tenure	< 3 Months	3 (5.1%)
	2- 12 Months	17 (28.8%)
	1-2 Years	10 (16.9%)
	2-5 Years	13 (22.0%)
	>5 Years	12 (20.3%)
Multi Focus	Yes	35 (59.3%)
High Response	1	12 (20.3%)
High Detection	1	14 (23.7%)
High Prevention	1	7 (11.9%)

Respondents from the charities made up the largest cluster group within the study (n=11, 22.0%), while the Welsh Government cluster group was the smallest cluster

observed, made up of 6 respondents (13%).<sup>26</sup> A fairly even distribution is seen in police regions with 34 respondents operating in a single region (57.6%) and 25 operating in two or more (42.4%). A majority of respondents are not government-based (n= 40, 67.8%) and received no funding from the EU transition fund (n= 42, 71.2%). Just over half of respondents have been in post for under two years (n=30, 50.8%). 59.3% (n=35) of respondents identified as ‘multi-focus’ agents. Tension detection (23.7%, n=14) was the highest priority action, followed by a response (20.3%, n=12) and prevention (11.9%, n=7).

#### ***4.3.1. Cooperation Frequency Clusters***

Seven different items were collected to measure the levels of perceived cooperation frequency between stakeholders in the CDN. Each item invited the respondents to indicate levels of cooperation frequency for each stakeholder group on a six-point Likert-scale<sup>27</sup>. Factor analysis was carried out using PCA with varimax orthogonal rotation. This helps reduce the data for ease of analysis in subsequent regression models but also provides initial evidence of stakeholder group clustering. Alongside the PCA, a PROXSCAL two-dimensional MDS solution was likewise run to provide an accompanying visualisation of clusters (Figure 13). The stress value for the MDS (0.03) indicates a robust goodness of fit, giving us confidence in high output validity. Table 9 shows the final rotated component output produced by the PCA of all seven clusters. An initial inspection of the table reveals seven stakeholder groups loaded into a meta-cluster, and therefore, none were excluded in later analysis. The three-component solution explained 77.48% of the variance, with component one contributing 34.29 %, component two 27.87% and component three 15.32%.

---

<sup>26</sup> However, this includes a combined response given by the central Welsh Government that represents an entire team that specialises in equality, community tensions and inclusion. Other respondents that feature in the Welsh Government respondents include Crown Prosecution Service stakeholders and government-based members of the ‘Hate Crime Criminal Justice Board’.

<sup>27</sup> Cluster group “self-ratings” during the PCA procedures were excluded from the analyses.

*Table 9: Factor loadings for cluster group cooperation frequency*

Item	Rotated Factor Loadings		
	Component 1	Component 2	Component 3
<i>How much do you cooperate with the following organisations for cohesion delivery in Wales?</i>			
Cohesion Coordinators	.921		
Cohesion Officers	.894		
The Police	.720		
Charities		.957	
Community Groups		.928	
Academics			.834
Welsh Government			.692

Cluster borders that reflect the components found in the PCA are applied to the MDS visualisation. The items that loaded into component one included: ‘cohesion coordinators’, ‘cohesion officers’ and ‘the police’. All three items loaded into the first component have a strong association with the regional implementation of cohesion agents. Hence, all three clusters have duplicate roles that operate regionally. The cluster proximities seen within the regional agent meta-cluster (component one) on the MDS plot provide further evidence of the existence of sub-clusters within meta-clusters. Although the entire cluster border reflects the high scoring variance loading score, a much smaller sub-cluster border shows significantly lower loading variance between two meta-cluster member groups (cohesion coordinators and cohesion officers). This finding indicates that while the police fall into the same meta-cluster (regional agents) as cohesion coordinators and cohesion officers, they can be considered an outlier within this cluster.



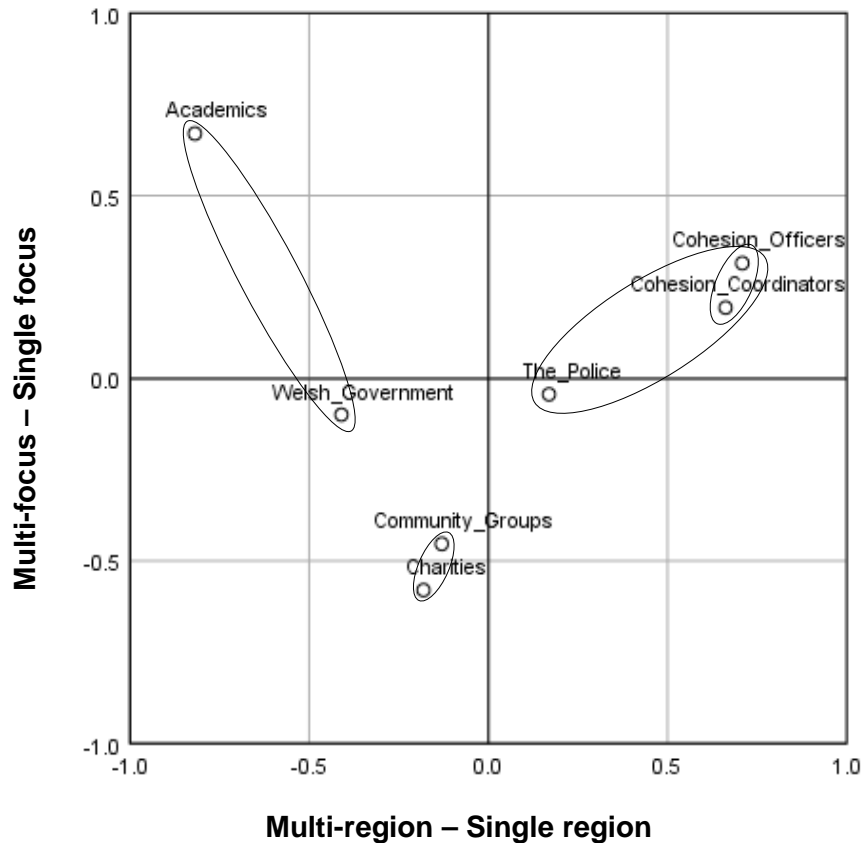


Figure 13. CDN cooperation frequency space

Component two loaded two items; ‘charities’ and ‘community groups’. Both clusters can be understood as non-governmental and have weak loadings with components one and three. The third sector meta-cluster (component two) shows a very small cluster border, meaning that cooperation frequency is high within component two stakeholder groups. This reflects the loading variance score that ranked first amongst the meta-clusters. The final component also loaded two items: ‘academics’ and the ‘Welsh Government’. These clusters have a strong association with policy and evidence. However, component three ranked third in loading variance. This is evident in the visualisation (Figure 13) that shows a significantly larger cluster border between academics and the Welsh Government. Labels were developed for each meta cluster. Component one: cooperation frequency with regionally implemented cooperation agents; component two: cooperation frequency with NGOs; and component three: cooperation frequency with policy and evidence agents.

Inter relationships refer to relationships between stakeholders in different meta-clusters, whereas intra relationships are seen between stakeholder groups in the same meta-cluster. These can be further understood by considering a rank order of mean stakeholder group cooperation frequency. Clear variability between cluster groups is evident on the plot. For example, academics can immediately be identified as outliers within the CDN. Figure 14 supports this, which shows an accompanying rank order of all node relationships. All of the six lowest ranking cooperation frequency relationships involved the academic group. Most of these cases (coordinator: 2, officer: 1.56, police: 2.11 charity: 2.46) indicated cooperation occurring every few months. The academic- Welsh Government relationship had a mean score of 2.75 (monthly cooperation). The relationship case with least frequent levels of cooperation throughout all nodes was seen between academics and community groups (1.48). Moreover, the academic cluster group has a mean rank of 18.50. This indicates very low levels of cooperation frequency and was ranked the lowest (7th) of any cluster group.

Conversely, the Welsh Government had the highest mean rank (7.67) amongst the cluster groups, ranking first. This was closely followed by the Cohesion Coordinator cluster group, the second highest mean rank (8.67). Community groups (mean rank: 10.5, overall rank: 4th) and charities (mean rank 11.67, overall rank 5th) can likewise be identified as cooperation outliers, although both of their mean ranks are significantly closer to other nodes than academics.

Community group and charity cluster groups reported high levels of cooperation with other clusters. For example, community groups experience high levels of cooperation with the Welsh Government (mean score of 3.6- weekly cooperation, rank: 6th) and cohesion officers (mean score of 3.31- monthly cooperation, rank 10th). This suggests that cohesion officers and Welsh Government are a gateway for community groups to enter the CDNs dense cooperation space. Despite the community group and charity cluster groups generally being outliers of the CDN, the relationship between them is very strong, ranking third in the entire network (mean score of 3.93 – weekly cooperation). The charity-community group relationship is therefore identified as a 'high-frequency relationship', this is expected, because they belong to the same meta-cluster (as determined by the PCA). High-frequency

relationships are defined in this study as any relationship that averages weekly or above. Another high-frequency relationship is evident on the MDS plot between cohesion coordinators and cohesion officers (mean score of 4.62 –daily cooperation). As previously discussed, this is a high cooperation sub-cluster within the regional agent meta-cluster.

### 4.3.2. Stakeholder relationship cooperation rankings

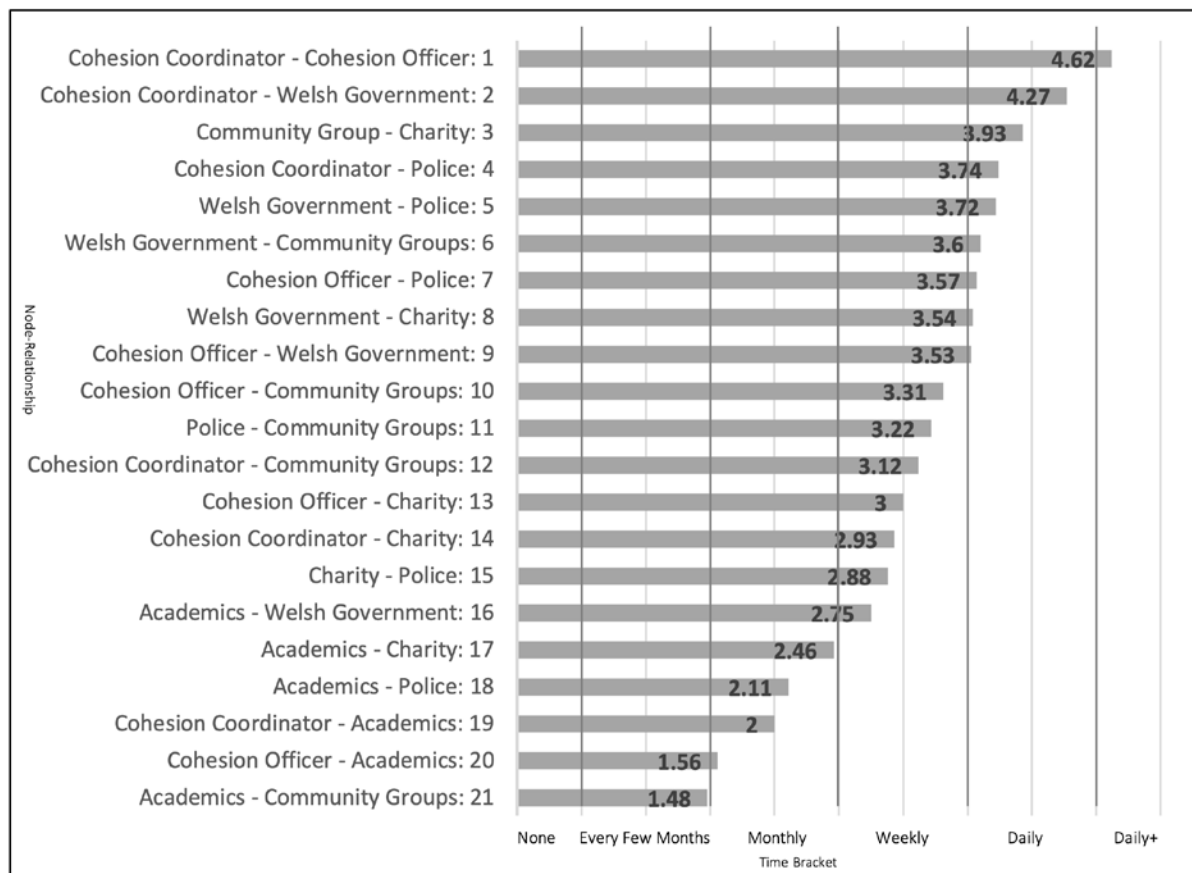


Figure 14: CDN node relationship rank order

A more detailed inspection of the rank table<sup>28</sup> reveals that an inner circle of cooperation exists between cohesion coordinators, cohesion officers, police, and Welsh Government. Despite not having the highest mean rank, cohesion coordinators experience some of the highest levels of cooperation with other inner

<sup>28</sup> Note: 'daily+' is several times a day

circle members, ranking 1st, 2nd, and 4th on the comparative table. The highest levels of cooperation, that do not involve cohesion coordinators, exist between the Welsh Government and the police (mean score of 3.72, rank: 5th). This finding, alongside the relationship observed between the Welsh Government and cohesion coordinators (mean score: 4.27, rank: 2nd), indicate that although the Welsh Government did not load onto the regional meta-cluster, it still enjoys high levels of cooperation. Conversely, the academic group shows low cooperation frequency with regional meta-cluster nodes. This suggests that the Welsh Government is a gateway to bridge academic nodes and regional agent (cohesion coordinators, officers, and police) nodes. The police saw frequent levels of cooperation with the cohesion coordinators (mean score of 3.74, rank: 4th) and cohesion officers (mean score of 3.57, rank 7th). These results help explain the nature of the regional agent meta-cluster.

Cohesion coordinators and cohesion officers have a clear 'cohesion team' sub-cluster, with the coordinators serving as the bridge between the police and cohesion officers in the wider meta-cluster. The mean cooperation frequency of all relationships surveyed is 3.11 (monthly cooperation). Concerning the PCA determined meta-clusters, the mean score for intra-cluster relationships is 3.72 (weekly cooperation), whereas the mean score for inter-cluster relationships is 2.92 (monthly cooperation). When breaking down intra-cluster mean scores, we can see that intra relationships between the regional agent member stakeholders (cohesion coordinators, cohesion officers and the police) have the highest cooperation frequency (3.98), followed closely by third sector (charities and community groups) intra relationships (3.93). Despite a higher average cooperation frequency for intra-cluster than inter-cluster relationships, one exception can be seen. The policy/evidence intra mean score is lower than inter relationships between; (1) third sector-policy/evidence and (2) regional-policy/evidence. The inter meta-cluster relationship with highest cooperation frequency is between regional agents and policy/evidence (2.87).

### **4.3.3. Cooperation MDS Axes**

Understanding that the MDS output shows a visualisation of cooperation space between cluster variables enables us to assume that the dimensions reflect particular characteristics about the nodes involved in the survey. A further examination of the plots concerning the x and y axes shows clear distinctions between cluster variables. On the y-axis, nodes that take on a generalised multi-focus (academics, cohesion coordinators, cohesion officers, Welsh Government, and the police) approach about differing protected characteristics dominate the higher end. Conversely nodes that take a one-focus approach (community groups and charities) dominate the lower end. The nature of how multi/ non-multi focus is defined is discussed later in this section.

On the x-axis, organisations with a network of different regional teams dominate the right side. This includes cohesion officers, cohesion coordinators and the police. Cohesion coordinators and officers operate in eight regions: West Gwent, East Gwent, Cardiff and the Vale, Cwm Taff, Swansea Bay, Mid and West, Northeast and Northwest. The police operate in four main regions: South Wales, North Wales, Dyfed Powys, and Gwent. However, there is further local division within each of the four regions. Organisations that involve one centralised team, that can still work across two or more regions, occupy the right side. This is supported by descriptive statistics from a question asking respondents to state which police regions they operate in. The output shows minimal cases of working across 'two or more' regions for; cohesion coordinators (0%), cohesion officers (12.5%), and the police (11.1%). Conversely, the output showed high levels of working across 'two or more regions' for cluster variables such as the Welsh Government (100%), community groups (40%) and academics (75%), on the left side of the axis.

### **4.3.4. Regression Models**

The models used two predictor sets: (1) cluster characteristics and (2) cluster perceptions. Cluster characteristic predictors relate to node's self-reported instances of high levels of tension detection/response/prevention, working across two or more

regions and having a multi-focus (see multi-focus predictor). The cluster perception predictors were reflective of the PCA loadings and measured quality of cooperation. Both sets of predictors were regressed onto the dependent variables that were ascertained from the PCA loading for each model. Model 1 relates to cooperation frequency for the third sector meta-cluster (community groups and charities), Model 2 relates to cooperation frequency for the regional meta-cluster (the police, cohesion coordinators and cohesion officers) and Model 3 relates to cooperation frequency for the policy and evidence meta-cluster (academics and the Welsh Government). The results are presented in Table 10.

#### ***4.3.5. Cluster Characteristics***

Despite being less predictive than cluster perceptions, variables within the cluster characteristic set emerged as having significant associations with cooperation frequency in all three models, holding all other factors constant. In the first model, relating to the third sector meta-cluster (charities and community groups) three characteristics were found to have significant associations with cooperation frequency. Respondents who reported high levels of focus on preventative work ( $\beta = 0.48$ ,  $p < 0.01$ ) in the CDN were likely to have a higher reported cooperation. Respondents who took on a multi-focus approach ( $\beta = -0.55$ ,  $p < 0.01$ ) in the CDN were less likely to have high cooperation frequency. Finally, in relation to the third sector meta-cluster respondents that worked in two or more regions ( $\beta = 0.38$ ,  $p < 0.01$ ) were more likely to report higher levels of cooperation. For the second model, relating to the policy and evidence meta-cluster (academics and Welsh Government), only one cluster characteristic factor emerged having a significant association to cooperation frequency. Respondents with high levels of tension detection ( $\beta = -0.24$ ,  $p < 0.10$ ) in their work associated with having lower reported levels of cooperation frequency, although this association narrowly approached conventional levels of significance. In the third model, relating to the regional agent meta-cluster (cohesion coordinators, cohesion officers and the police), two cluster characteristics were found to have significant associations with cooperation frequency. Respondents who took on a multi-focus approach ( $\beta = 0.13$ ,  $p < 0.10$ ) were more likely to have a higher reported cooperation frequency. However, this

association also only just approached conventional levels of significance.

Respondents who have received funding from the EU transition fund ( $\beta = 0.14$ ,  $p < 0.05$ ) were found to have a significant positive association with cooperation frequency.

#### **4.3.6. Cluster Perceptions**

Initially, 'cluster importance' was included in the regression models in the cluster perception set alongside quality of cooperation. However, high levels correlation causing issues of multicollinearity indicated that cooperation quality and cooperation importance were not independent from one another and therefore could not both be included in the analysis. Thereafter, cluster importance was removed from all three models, leaving cluster quality (of cooperation). This question was applied to seven separate items for each of the cluster groups. The scores given by respondents are therefore more subjective than those given for quantity of cooperation, due to no metric categories being provided. The respondents instead indicate their perceived levels of cooperation quality based on interactions with other nodes in the CDN. Reliability analyses were conducted in order to ensure that data reduction for the same meta-clusters (third sector, regional agent, and policy/evidence) could be applied as three separate continuous predictor variables in the regression models.

Holding all other factors constant, significant associations were found between perceived quality of cooperation and cooperation frequency in all three models. In the first model, relating to the third sector meta-cluster (community groups and charities), two cluster perception predictors were significantly associated with cooperation frequency. A positive association was found between cooperation frequency (model 1) and the third sector quality of cooperation scale ( $\beta = 0.68$ ,  $p < 0.01$ ). On the other hand, a negative association was found between cooperation frequency (model 1) and policy/ evidence cooperation quality ( $\beta = -0.35$ ,  $p < 0.01$ ). The second model, relating to the policy/ evidence meta-cluster (academics and Welsh Government), showed that a positive association exists between cooperation frequency and policy/ evidence quality of cooperation ( $\beta = 0.61$ ,  $p < 0.01$ ). Finally, the third model, relating to the regional agent meta-cluster, showed two significant

associations. A positive association is observed between cooperation frequency and regional agent quality of cooperation ( $\beta = 0.77$ ,  $p < 0.01$ ). Conversely, a negative association was found between cooperation frequency and cooperation quality for the policy/evidence meta-cluster ( $\beta = -0.17$ ,  $p < 0.05$ ). These findings are further discussed later in this chapter.

*Table 10: BCa bootstrap OLS regression predicting CDN network cooperation*

	Third-sector Meta Cluster			Policy/ Evidence Meta Cluster			Regional Agent Meta Cluster		
	B	1 SE	b	B	2 SE	b	B	3 SE	b
<i>Cluster Characteristics</i>									
Detection	-0.47	0.29	-0.25	-0.54	0.30	-0.24*	0.05	0.15	0.02
Prevention	1.23	0.53	0.48***	-0.01	0.45	-0.00	-0.00	0.29	-0.00
Response	-0.13	0.44	-0.06	-0.06	0.42	-0.03	0.14	0.19	0.06
Multi-Focus	-0.93	0.27	-0.55***	0.10	0.30	0.05	0.26	0.15	0.13*
Multi-Region	0.62	0.19	0.38***	-0.03	0.30	-0.02	0.14	0.18	0.07
EU Transition Fund	0.05	0.33	0.03	0.13	0.33	0.06	0.34	0.17	0.14**
Government-based	0.25	0.29	0.13	0.47	0.30	0.20	0.10	0.14	0.04
<i>Cluster Perceptions</i>									
Regional Agents (Quality)	-0.02	0.16	0.15	-0.29	0.20	-0.30	0.79	0.10	0.77***
Third sector (Quality)	0.58	0.11	0.68***	0.20	0.16	0.20	-0.02	0.06	-0.02
Policy/ Evidence Agents (Quality)	-0.30	0.09	-0.35***	0.60	0.14	0.61***	-0.17	0.07	-0.17**
Constant	0.25	0.21		-0.03	0.26		-0.43	0.20	
<i>Model Fit</i>									
Sig	0.00			0.02			0.00		
R <sup>2</sup>	0.73			0.23			0.82		
n	59			59			59		

#### **4.3.7. Sub-Model Analysis**

Before examining individual variables, a sub-factor analysis can shed light on which set of variables (cluster characteristics and cluster perceptions) can best predict variance within each model. The adjusted R<sup>2</sup> metric was used for sub-model analysis. For the third sector meta model (model 1), cluster characteristic predictors accounted for 2 percent (R<sup>2</sup> .02) of variance, with cluster perceptions accounting for 46 percent (R<sup>2</sup> .46) of variance. In the policy/ evidence meta-cluster (model 2), cluster characteristic predictors accounted for 2.5 percent (R<sup>2</sup> .025) of variance, with cluster perceptions accounting for 30 percent (R<sup>2</sup> .30) of variance. Finally, in the regional agent meta-cluster (model 3), cluster characteristics accounted for 43



percent ( $R^2 .43$ ) with cluster perceptions accounting for 76 percent ( $R^2 .76$ ) of variance. To summarise, the sub-model analyses in all three models suggest that cluster perceptions are more predictive of cooperation frequency than cluster characteristics. Within both sets of predictors, a contrasting range of individual variables were significantly associated with cooperation frequency across all models.

#### 4.4. Discussion

In support of H1, this chapter has shown clear evidence of nodal governance being implemented for cohesion delivery in Wales (Shearing, 2001). Although the MDS and PCA procedures are helpful in both confirming all stakeholder groups play a part in cohesion delivery and providing insights as to which partners cluster with one another, they say little about why this is the case, and how collaboration is not being maximised. Thereafter, using open question output to build on the quantitative findings is important in defining what kind of nodal governance exists, and what can be adapted to improve it. MDS and comparative rankings indicate the Welsh Governments are central nodes to the network, exhibiting high levels of cooperation frequency. Lending further evidence to H1, other stakeholder groups alluded to the idea that the Welsh Government is primarily responsible for ‘agenda-setting’ and giving steer to network priorities. Although it is common for policy actors to be responsible for strategy direction in multi-agency partnerships (Betts, 2002), the way in which actors above, below and beyond the state are incorporated can alter significantly. The Welsh Government provided an overview for the CDN ‘Community Cohesion Programme’:

*“The Welsh Government takes a multi-agency approach to community cohesion in Wales. Whilst we fund the delivery of the Community Cohesion Programme, our policy is achieved through partnership working with many agencies, including the police, local authorities, and the third sector. Cooperation and joint delivery is essential in ensuring that cohesion is considered and promoted in a strategic way across all levels of Wales and directly within communities, as well as monitoring and addressing tensions.”-Welsh Government unified response*

The cohesion agenda shows clear intentions to incorporate a plethora of actors, forming a multi-agency approach (Burris, 2004). Additionally, the comments made in relation to communities resonate with the 'partnership' step of Arnstein's (1969) ladder of engagement. The partnership step suggests a degree of citizen power, in which community groups can veto or promote meaningful network and agenda propositions that significantly affect the overall strategy employed. While this is the apparent intention of the programme, the quantitative output indicates that community groups are not as central to the network as perhaps intended. One respondent suggested that there is currently not enough community participation in the network:

*"I think that there is more we could all be doing (for community engagement), partnership working is essential. It is so important that we have a good relationship with the public, so cohesion work is vital in all communities. We need to be able to know what is going on in the community to effectively police it. The Senedd will inevitably link in with HQ staff rather than divisions"- **Police hate crime officer***

The respondent explains that many issues and tensions within communities can be understood and visible, if frequent and effective contact is in place. They suggest this is ascertained by creating good relationships, with strong rapport. They suggest that policymakers perhaps prioritise HQ staff over specific divisions. They indicate that taking on a more centralised approach can harm tension monitoring within communities due to less direct public consultations. This resonates with findings made by Thomas (2010), in which state agencies, such as the police, exhibited 'turf expansion' that diminished the involvement and achievement of objectives of other stakeholders such as local authorities and community groups. This sentiment is further reinforced from a community group respondent:

*"I'm sure other communities have had a lot of joy with the Welsh Government and other partners, but sometimes we feel a little left in the dark. I haven't even heard of "cohesion officers"- **Community group respondent 8***

Comments from these respondents, coupled with the output, indicate that some of the issues found by Thomas (2010) in the PVE also exist in the CDN. Community

engagement in the CDN is currently characterised by a 'degree of tokenism' rather than a 'degree of citizen power' (Arnstein, 1969: 217) suggesting that denizenship (Shearing and Wood, 2003) may be absent in a meaningful way within the CDN. The current involvement of community groups reflects the 'placation' step, wherein citizens are given limited degrees of influence. This indicates a need for greater emphasis on active citizen participation and engagement strategies that foster denizenship, enabling marginalised voices to be effectively heard and integrated within multi-agency systems (Blake et al., 2008). As previous research has shown (Chakraborti, 2018; Garland et al., 2022), the lack of meaningful integration can pose a significant threat to the ability of networks to detect, respond to, and prevent incidents, crimes, and underlying tensions. This fragmentation leads to reduced reporting of incidents and may worsen such issues. This lack of integration can result in duplication of efforts, miscommunication, and a lack of coordination among network members. This can ultimately harm the community, as incidents and underlying tensions may go undetected or unaddressed, leaving a significant dark figure in police reporting (Schweppe et al., 2020). Although current community involvement may not be as significant as intended in the cohesion programme outline, this could be attributed to the relatively recent development of the community engagement specific roles, such as the cohesion teams. Although the cohesion teams and community groups did not load into the same clusters, some evidence suggests that they will further integrate communities into the CDN when their roles are more developed. Respondents indicated that cohesion officers were an important group to help involve communities within the network:

*"Cohesion Officers are on the ground within communities and act as a bridge between their organisations and communities" - **Police policy officer***

The extracts indicate that the coordinator acts as a bridge between CDN organisations and communities, and the officer's role is predominately 'on the ground' working directly with communities. This finding indicates that the CDN differs from the PVE. Thomas (2010) found that the police overreach into local authority responsibilities and hinder their ability to interact with community groups. However, in the case of the CDN, evidence from clustering and the extract above indicate that

the police regard local authority departments such as the cohesion teams as a valuable part of the network, that can act as a strategic bridge to link the overall key agendas, devised by the Welsh Government (Betts, 2002). Therefore, inter meta-cluster relationships between community groups and regional agents are integral in helping to integrate community groups into the CDN 'inner circle' thus linking overarching policies and agendas with those most effected by them. With this in mind, it is clear that as these relationships are strengthened a notable shift from the current state of placation to the initially intended degree of citizen power could occur. A reason for community groups being more likely to interact with cohesion teams was suggested by a charity respondent:

*"Our experience of working with community groups is that they feel more comfortable plugging into specialised roles (cohesion teams) than other partners" - **Charity volunteer***

The extracts offer evidence supporting the effectiveness and engagement of specialised roles in addressing specific issues within communities, potentially due to their informal nature. This indicates that cohesion-specific roles could possess higher levels of symbolic capital than other partners, including the police. Communities perceive such roles as more trustworthy and effective in addressing core issues. This lends further evidence to the idea of integrating third party organisations into the reporting process to enhance overall accuracy (Schweppe et al., 2020). The effectiveness of these roles is supported by quantitative findings that demonstrate a higher quality of cooperation between communities and cohesion teams than with the police. As earlier extracts also suggest, there are reporting issues between the police and communities, as admitted by a police stakeholder. The lack of trust and social capital between affected communities and police can lead to underreporting of incidents in rural communities. This raises concerns about the deficits in symbolic capital, which has been observed in similar communities (Chakraborti & Garland, 2003). Although the involvement of the cohesion team shows clear benefits to the network, further extracts show reasons why they are not as central to the cooperation space as other nodes like the Welsh Government.

*"I didn't know cohesion coordinators and officers existed unless they are in touch through a different title"- **Charity director***

*“I am not used to this term 'Cohesion Coordinators' / 'Officers' on an equal footing with such established groups as the police and WG”-  
**Community group respondent***

Evidently, many other stakeholders see the value in both cohesion roles. However, many partners do not interact with the cohesion teams or in many cases are unaware of their existence. The second extract suggests that the coordinators and officers aren't put onto 'equal footing' within the CDN as other cluster groups such as the Welsh Government and the police. The cooperation data and feedback relating to the cohesion teams indicate a high quality of cooperation with other network stakeholders and communities. As such, cohesion teams are well-positioned to build strong social capital with communities, resulting from their high levels of symbolic capital (Dupont, 2004).

Consequently, they have the potential to be effective gateways for community groups to enter the community development network, which could enable citizens to operate with a degree of power in community justice. However, despite this potential, many communities that could benefit from the cohesion team's input are unaware of their existence. This finding reinforces previous research that suggests poor integration of communities into multi-agency networks can lead to limited knowledge of certain policies, support services, or agencies that are intended to assist, but are instead overlooked due to a lack of awareness (Hardy and Chakraborti, 2020; Garland et al., 2022).

This lack of awareness has prevented the optimisation of strong social capital building between communities and cohesion teams. These highlights significant gaps in the network. Cases of community groups being unaware of these roles can perhaps be attributed to the length of time officers have been in post. Descriptive statistics show that none of the cohesion officer respondents have been in post for over two years. Most cohesion officer respondents have been in post between 3 and 12 months (87.5%, n=7), with the remaining respondents having been in post between 1 and 2 years (12.5%, n=1). Therefore, carefully considering cohesion roles as they are further developed, recognised, and established is advised to bridge the gap between traditional agencies and communities. In doing this, the network can

ensure a meaningful and non-tokenistic involvement of communities, whom cohesion delivery concerns the most. Alongside, current difficulties to fully integrate communities into cohesion delivery, issues of overall CDN agendas lacking a cohesive framework were suggested in the open output:

*“WG needs to have a greater visibility and transparency over its leadership and strategy and coordinate a range of statutory and non-statutory bodies and community groups- **Police policy officer***

Past studies have shown that when utilising a pluralised approach, complications arising from differences in organisational cultures and structures mean that collaboration, communication, and transparency are key principles required from corresponding agencies to facilitate effective working partnerships (Kean and Hamilton, 2004). However, many respondents suggest that the network lacks a cohesive agenda for all stakeholders to follow due to minimal clarity. This reflects Zedner’s (2007) criticisms of an expansive and potentially intrusive state. Although in this case state objectives are not actively overshadowing other sector priorities, the lack of a clear agenda perhaps hinders how effectively they can be addressed. The Welsh government holds significant political capital (Dupont, 2004) in the network, as they possess considerable capabilities to influence or even determine policy. As a result, the responsibility for defining and coordinating network strategies falls upon them. The respondents therefore recommend that the Welsh Government provide a clearer ‘vision’ and ‘definition’ of community cohesion within the policy realm to all current stakeholders (Betts, 2002). Implementing a more cohesive agenda, could likewise aid in improving community participation, as one cohesion coordinator suggests:

*“There also needs to be a clearer vision and definition of community cohesion. Welsh Government needs to have a greater visibility and transparency over its leadership and strategy, they should coordinate a range of bodies and community groups. They should be acting as the main points of contact for cohesion projects, raising awareness and collaboration between all groups”- **Cohesion coordinator 5***

Furthermore, other partners suggest that the Welsh Government should implement a reimagined agenda, which could ‘improve cooperation’, ‘reduce duplication faults’

and provide an opportunity to better integrate communities into the overall framework at a local level. Results from the PCA and MDS analysis gave evidence to support that stakeholder clustering exists in the CDN. In total, three main stakeholder clusters were found: (1) regional agents, (2) third sector agents and (3) policy/evidence agents. No stakeholder group was entirely alienated from the network, with all loading into a meta-cluster. The dimensions provided on the plot gave an insight into why some of these clusters formed. Two main attributes associated with the plot axes and therefore attributed to stakeholder clustering: (1) agency regionality and (2) multi-focus prioritisation. Multi-regionality refers to partners with duplicate roles in various regions to provide local coverage. The regional agents cluster involved all stakeholder groups that operate under multi-regionality. Open questions supported this, with respondents indicating that they find cooperation easier with partners in their local area:

*“I personally find the cohesion teams helpful for my work because they assist with the day-to-day local issues”- **Police hate crime office***

Moreover, using localised teams to address more specific issues is beneficial. That being said, some issues were identified between regional partners. The two lowest ranking intra meta-cluster quality relationships are seen between the police and cohesion coordinators (variance: 0.79, rank: 7th) / officers (variance: 1.13, rank: 9th). Potential reasons for this were given in open questions by members of the cohesion team:

*“Often some hate incidents that don’t fall above the hate crime threshold aren’t shared with us by the police, we always try and pass on as much information as possible to them”- **Cohesion coordinator 1***

This indicates a collaboration gap in the network, resulting, in part, from non-reciprocal communication. This is consistent with ideas put forward by Cheminais (2009) that multi-agency partnerships can be weakened by a poor culture of information-sharing, resulting in a drop in cooperation frequency. Although regional stakeholders such as the police and the cohesion teams suggest the existence of localised partners can be highly beneficial, issues such as information-sharing can

still arise. This is perhaps a product of a lack of clarity in the overall multi-agency agenda the Welsh Government sets. When re-considering the CDN framework, indicators for stakeholders of what, how and when to share information with other partners could reduce duplication faults common to multi-agency networks (Cheminais, 2009) enhancing cooperation and productivity.

Multi-focus prioritisation can likewise be attributed to stakeholder clustering. Both regional-agent and policy/ evidence clusters take on a multi-focus approach. Conversely, the third sector cluster (community groups and charities) nodes tend to focus on only one protected characteristic cluster. This was reflected in the open question output:

*“For us unless it’s a major or criminal issue we tend to only really interact with charities. because they have a greater focus on issues that matter to us”- **Community group respondent 2***

Indeed, descriptive statistics, likewise showed that third sector stakeholder groups were predominately made up of single-focus nodes. These findings resonate with Suhr’s (2005) suggestion that stakeholder groups with agenda commonalties and homogenised focus areas are more likely to cluster in multi-agency partnerships and build strong social capital (Dupont, 2004). Although, this may be true, it perhaps can lead to alienation of particular groups, particularly within communities and may also be a source of lower community participation in the CDN. The only stakeholder group that exhibits lower cooperation in the CDN is Academics. However, this may not necessarily be negative, a closer examination of open output and the policy/evidence meta cluster loadings can help further explain their role in community justice networks.

The policy/evidence meta-cluster had the highest PCA variance, indicating low overall cooperation. When examining individual cross-node relationships, we can see that although both belong to the same meta-cluster, the Welsh Government has higher cooperation frequencies with all other non-academic nodes. Therefore, the Welsh Government and academics loading into the same component more supports the idea that the Welsh Government serve as a gateway for academics to enter the



CDN. While limited academic collaboration could potentially pose challenges, echoing Chakraborti's (2016:1) description of it as "tokenistic" or a mere checkbox exercise within multi-agency approaches, respondents suggested that despite infrequent cooperation, scholarly contributions to the Welsh CDN might still be meaningful:

*"Personally, I don't know if more face time is even needed with academics, so long as their research is relevant to us and helps give evidence led approaches"* - **Charity director**

*"I only really ever plug in through the Welsh Government who distribute"*  
- **Academic**

The extracts allude that academics don't need to be a day-to-day feature of the CDN, but instead a tool for helping devise evidence led approaches. It shows that their input is valuable, with a high degree of informational cultural capital (Dupont, 2004). However, their input is not dependent on how frequently they cooperate. Instead, cooperative efforts need to be integrated at the right time to produce valuable outcomes. This idea is further supported by findings in the regression model, indicating that greater cooperation quality between Academics and other partners was negatively associated with increased cooperation frequency. In other words, better input made by policy/ evidence nodes reduces the necessity for cooperation with other stakeholders. Open output suggests this could be true for not only academics but also the Welsh Government:

*We only get involved with some of these groups (Welsh government etc) when things go wrong. So I'm not sure if more is even always better. The overall goal for me would be to eradicate hate, so approaching them wouldn't even be necessary* - **Charity executive director**

*My role doesn't really have much direct contact with the WG. I'm not sure increasing it would even help too much-* **Cohesion officer 3**

The relationship between cooperation quality and frequency was further tested amongst all stakeholder groups. Findings in the sub-model analyses for all three

models supported H3. They were consistent with past research findings showing node perceptions are more predictive than any other sub-factor (Levi and Williams, 2013). However, unlike the policy/ evidence cluster, the association between cooperation frequency and quality was positive in every other case, including internally between academics and the Welsh Government. This indicates that improved communication and information-sharing quality is important in enhancing network productivity.

Less evidence existed to support the fourth hypothesis that stakeholder role characteristics would predict cooperation frequency. Cluster characteristics included multi-focus, multi-region, government-based, EU transition fund recipient and tension agendas (high focus: detection, prevention, and response). First, cluster characteristics explained predicted cooperation less than cluster perceptions in all three model's sub-factor analyses. This was particularly true in the findings for both the policy/ evidence and regional agent meta-cluster models. The policy/ evidence only had one association with tension detection and cooperation frequency that approached conventional significance levels ( $<0.1$ ). The regional agent cooperation frequency meta-cluster had no significant associations with the three tension agenda predictors. However, EU transition fund recipients were likelier to have high cooperation frequency with regional agents. Interestingly, most funding recipients belong to the regional agent meta-cluster, particularly within the cohesion teams. This supports previous suggestions that shared agendas from funding can result in stakeholder clustering and increase cooperation frequency (Kean and Hamilton, 2004).

More significant associations were found in the NGO cluster. Despite this, they still accounted for low levels of variance in the sub-model analysis. Only one tension agenda (prevention) was positively associated with cooperation frequency about third sector cooperation. This finding suggests that prevention-based agendas are important amongst third sector partners and can lead to increased cooperation. Multi-focus agendas (concerning protected characteristics) were negatively associated with cooperation frequency with third sector partners. This reflects descriptive statistics that showed that only two stakeholder groups with low levels of multi-focus belonged to the third sector meta-cluster. This is consistent with past

multi-agency literature that expectedly shows differences in stakeholder agendas reduce cooperation frequency (Cheminais, 2009). Moreover, although less evidence suggests that node characteristics significantly predict cooperation frequency more than cluster perceptions, some factors have high associations. Another consideration is that other node characteristics not studied in the models may explain greater variance.

The COVID-19 pandemic has notably impacted the CDN, affecting how organisations and communities collaborate. Despite initial fears that the pandemic would negatively affect cooperation between agencies and partners, it has led to more frequent inter-agency cooperation. Real-time communication (RTC) platforms, such as Microsoft Teams, became more prevalent during the pandemic, allowing for easier virtual meetings and better coordination between organisations. As a respondent pointed out:

*"People have become less insular and more willing to share ideas/best practice. Meeting more regularly via Microsoft Teams." – Cohesion coordinator 5*

This has resulted in improved engagement and liaison with partners and less travel time, allowing for more time to work and interact with other stakeholders. As another respondent noted:

*"More frequent meetings with stakeholders leading to more frequent exchanging of information." – Hate crime manager*

However, while virtual meetings have been beneficial in many ways, they have also created challenges for some organisations, particularly those involved in areas such as victim support, where face-to-face interactions are crucial. As a respondent noted:

*"We were told that Zoom is no longer used because of privacy issues and zoom bombing? It's so hard to keep up."- Charity officer*

*"Poorer in some areas as more vulnerable groups are harder to reach."  
-Cohesion coordinator 3*

*“The Hate Crime Criminal Justice Board meetings were predominately in Cardiff; now they are always virtual, it feels less Cardiff-centric”-  
**HCCJ Board Member***

This indicates that services that require on-the-ground attention have been negatively affected. Additionally, the lack of a standardised RTC platform has created additional challenges for some organisations. While the inner circle of the CDN, such as the Welsh Government and police, have opted for Microsoft Teams, other community groups and charities may use different platforms, such as Zoom, making arranging online discussions difficult. Despite these challenges, many organisations have found ways to work more closely with each other during the pandemic. Virtual meetings have facilitated better communication and information exchange, improving partner engagement and liaison. Moreover, the use of virtual meetings has led to a more equal footing not only amongst agencies but also between regions. Furthermore, the COVID-19 pandemic has significantly impacted the CDN, creating challenges and opportunities for organisations and communities. While virtual meetings have facilitated better communication and coordination between organisations, the lack of face-to-face interactions has negatively impacted citizen engagement, particularly for vulnerable groups. Nonetheless, organisations have found ways to adapt to these challenges and work more closely together, which will likely have a positive impact in the long term. As one respondent noted:

*"An improvement in engagement and liaison with partners which should be an important legacy of the pandemic." – **Police liaison unit***

Some survey questions aimed to explore the integration of emerging methods of online tension monitoring. In this thesis the term online tension monitoring more specifically relates to social media channels and is defined systematic analysis of social media data from platforms like Twitter, Facebook, and YouTube to detect signs of heightened societal stress or potential conflicts. It enables real-time observation of social media streams to identify deviations from normal social cohesion, aiding in early detection and response to emerging tensions. This approach leverages computational tools to analyse user-generated content, offering insights that complement traditional methods of community intelligence gathering (Burnap et al., 2015).

Specific questions were designed to elicit respondents' opinions on implementing online tension monitoring methods and understand ideas being used offline. One question, in particular, focused on social media and asked respondents to indicate their agreement with the idea that further methods of social media monitoring should be implemented by stakeholders to understand underlying tensions<sup>29</sup>. Analysis of the responses revealed that most respondents, 82.5%, believed that further methods of online tension monitoring should be implemented, with 37% and 45.7% agreeing "a great deal" and "a lot," respectively. The breakdown of responses by cluster group revealed that most respondents across all groups agreed to a high level, with the police and community coordinators being the highest at 89.9% and 87.5%, respectively. These findings suggest a strong endorsement for online tension monitoring methods across the network. To gain a deeper understanding of what potential tools should entail, further questions focused on identifying why a standardised dashboard tool would be helpful and the requirements needed from any potential tools. Stakeholders believe that the current manual processes for monitoring tension in communities are time-consuming and inefficient, and have expressed a need for an online tension monitoring dashboard that could help capture and analyse social media data quickly and effectively:

*"Currently, we are not good at collecting tension, which is online. We need to have an automated mechanism to assist us with capturing all relevant online tension asap."* - **Cohesion coordinator 1**

*"It would be more cost, time and resource-effective to implement an online tension monitoring dashboard"* - **Academic**

These suggests that an automated tension monitoring dashboard could be instrumental in reducing the time and resources currently devoted to manual monitoring processes. In this context, stakeholders believe such a tool would provide real-time analysis, allowing for quicker responses to emerging tensions, thereby freeing up resources and time for other activities. Stakeholders also indicated that social media is an effective method of gathering data, of tensions before they

---

<sup>29</sup> This item was measured using a Likert scale with five response options ranging from "none at all" to "a great deal."

escalate. This could aid in taking a preventative approach towards community tensions by addressing issues before escalation:

*"You need to monitor well so that you can respond appropriately and tackle problems before they grow." - **Charity director***

*"General non-specific hate speech is not captured which would help us identify where our Prevent work should be directed."- **Cohesion coordinator 7***

The second excerpt outlines additional reasons why stakeholders support the implementation of a tension monitoring dashboard, namely that it would be an effective means of identifying hidden tensions that are not presently detected. One issue for this, is the current reactionary framework for online tensions in place for organisations such as the police:

*"At present, as a force we are unable to capture online hate speech, unless it is reported to us. Usually that happens where there is a victim. General non-specific hate speech is not captured which would help us identify where our Prevent work should be directed." - **Police policy officer***

The framework in place for documenting online tensions is reactionary and has proven ineffective in capturing unreported tensions. This inadequacy becomes particularly concerning as stakeholders have themselves acknowledged that communities are reluctant to report such issues to authorities such as the police. This shortcoming reflects a symbolic capital deficit as described by Dupont (2004), where communities do not view the police as a reliable or effective institution for addressing these matters, leading to a depletion in social capital and the subsequent breakdown of communication between the groups. The broader implications of this network fault are explored in greater depth in the discussion section.

*"Many do not feel comfortable reporting issues to the police, these incidents and issues then fail to be recorded which leads to a misrepresentation of the situation."- **Charity coordinator***

Respondents advocated for a real-time tension monitoring dashboard to curtail such issues, rather than relying solely on reported incidents to capture and analyse all online tensions. This "fly on the wall" approach would allow for a more covert monitoring of genuine tensions, as perpetrators would not be conscious of being monitored. Additionally, the dashboard would enable the aggregation of larger trends, facilitating the prevention of widespread tension issues.

*"Many tensions exist online that are not identified. Understanding these will help prevent tensions at face value, and also on a macro level." – Academic*

To gain insight into the requirements of a successful dashboard for monitoring inter-group tensions in Wales, follow-up discussions were made with cohesion coordinators responsible for monitoring such tensions. They revealed that the previous system used for monitoring, CAT-D, was discontinued due to its high cost, and the coordinators expressed dissatisfaction with its reliance on Twitter data, seen as unrepresentative. Moreover, the coordinators highlighted the importance of a replacement system, as the current manual methods of observing tensions on social media are resource-intensive and insufficient. The study identified several key requirements for a future monitoring tool. Firstly, the system must be cost-effective while providing outcomes similar to CAT-D. Secondly, the system must be able to geo-tag tensions hotspots, as the coordinators require specific data relating to their region to develop a strategy in addressing the tensions. Thirdly, the system must not have biases towards certain demographics, and other publicly accessible social media platforms can provide more findings that apply to wider social groups. Fourthly, the system must monitor online tensions in an easy-to-use and accessible way, as the current manual forms of observing tensions on social media are ineffective and inefficient. Finally, the system must provide intelligence on resulting offline tensions, as offline tensions can often result from online interactions, and monitoring both can provide a more comprehensive understanding of the situation.

One potential tool for monitoring tensions is the Hate Lab dashboard. One contributor to the Hate Lab emphasised the Platform's potential as a monitoring tool for group tensions. The respondent explained that the Hate Lab dashboard is

developed specifically for monitoring and analysing digital tensions and other forms of online abuse across various social media platforms. In particular, it filters antagonistic speech related by race, religion, nationality, migrant/refugee status, disability, transgender identity, and sexual orientation. Compared to the discontinued CAT-D system, the dashboard provides access to additional platforms, such as Reddit, 4chan, and Telegram (with more to be added), which provides a wider diversity of social media data. It also offers an interactive and user-friendly interface, which enables quick and easy aggregation of complex data and its visualisation, allowing for an inspection of tensions over time and place<sup>30</sup>. The respondent also highlighted the potential to identify patterns of online inter-group tension that relate to offline phenomena and make informed decisions about strategies for combating them. For example, it facilitates targeted counter-speech into online networks and address issues related to inter-group tensions. Overall, the HateLab Platform<sup>31</sup> has the potential to provide a cost-effective and useful monitoring tool for inter-group tensions in Wales, which addresses the key requirements highlighted by the cohesion coordinators.

#### **4.5. Summary**

This chapter has provided an exploratory account of Wales's multi-agency partnership responsible for cohesion delivery. Although overall prevalence is not suggested due to a non-probability sampling method, initial evidence is given for inter-relationships between key CDN stakeholders. The mapping of the CDN is possible using MDS procedures. The visualisations reflect further information shown in PCA component loadings, and rank tables. By cross-examining all available information, the locations of gaps and bottlenecks in the CDN were identified. Further insights into the reasons for these gaps and bottlenecks were obtained through qualitative accounts obtained from participating stakeholders. One

---

<sup>30</sup> To address the need for geo-locating content, the respondent stated that limitations imposed by geo-tags (GPS) means alternative methods need to be used, including using place names, and issues relating to specific regions.

<sup>31</sup> The platform has been successfully piloted by a number of other organisations (see Annexes 1 and 2)



reoccurring factor is the lack of a coherent and cohesive plan for all partners to follow. This was found to weaken coordination, perhaps, causing cooperation faults, such as duplication gaps and poor information-sharing. Additionally, evidence shows the potential existence of a tokenistic involvement of community groups in the CDN. This more reflects Arnstein's (1969) steps of placation and consultation, rather than the apparently desired levels of meaningful engagement set out by policy partners. Evidence from the asymmetric quality of cooperation variance scores shows that non-reciprocated cooperation has a higher prevalence in inter meta-cluster relationships than intra meta cluster relationships.

However, in some cases intra relationships experience high levels of variance due to factors such as poor information-sharing. This results in sub-clusters, as seen in the regional agent's cluster, that can result in alienation of partners such as the cohesion coordinators. This however cannot be attributed to notions of turf expansion as seen in other networks (Thomas, 2010) but instead a lack of clarity between partners in the network. In fact, evidence suggests that turf expansion and empire building are far less evident between the police and local authority cohesion teams than in other networks such as PVE, with police stakeholders suggesting that cohesion teams are integral for bridging communities with the wider network. Community groups that have interacted with cohesion teams further supported this in the free-text fields. However, this effect has been minimised by many communities being unaware of the cohesion teams, thus reducing the extent of citizen engagement. This perhaps alludes to the idea that the degree of citizen participation (Arnstein, 1969) could naturally be enhanced in time as more community groups become aware of the cohesion team's existence. Gateway dynamics were likewise seen in the relationship shared between the outlier group - academics and the most central node in the inner circle- the Welsh Government. To summarise, this chapter provides evidence of the cooperation space in the CDN and indicate how and why some gaps exist.

Although many issues, such as stakeholders being unaware of other agencies existence, duplication faults and poor information-sharing exist, and can broadly be attributed to (1) the absence of a clear universal agenda shared between all partners including communities and (2) the relatively short amount of time some roles have been in place. This study recommends that partners within the CDN and similar

networks should be aware of these and be in open discussion when moving forward to curtail and minimise further development. The study also investigated the implementation of online tension monitoring methods in Wales. Most respondents supported this idea, and stakeholders expressed a need for a cost-effective and unbiased dashboard tool that can capture all relevant online tensions. The current manual processes are time-consuming and inefficient, and a monitoring dashboard can provide real-time analysis and aid in taking a preventative approach towards community tensions.

Additionally, several key requirements were identified for a future monitoring tool, including cost-effectiveness, geo-tagging of tension hotspots, absence of demographic biases, easy accessibility, and monitoring of resulting offline tensions. Finally, the research explored how the COVID-19 pandemic impacted the CDN. Virtual meetings and real-time communication platforms like Microsoft Teams facilitated better collaboration and coordination between organisations, resulting in improved engagement and liaison with partners. However, the lack of face-to-face interactions negatively affected citizen engagement, particularly for vulnerable groups. Nonetheless, organisations found ways to adapt to these challenges, leading to more effective and efficient working in the long term. The legacy of this pandemic could be improved collaboration and communication among organisations, which will likely have positive impacts on the CDN in the future.

## Chapter 5

### Exploring the Nexus Online and Offline Tensions: A Case Study of the Penally Asylum Accommodation

---

#### 5.1. Introduction

The announcement of the Penally asylum accommodation on September 14th, 2020, sparked controversy, and community tensions. The camps were intended to provide temporary accommodation for refugees seeking asylum in the UK, primarily from war-torn countries such as Syria and Afghanistan, while their asylum claims were processed. However, concerns regarding the suitability of the rural location and inadequate infrastructure, compounded by the perceived insufficient involvement of the local community in the decision-making process arose after its announcement. Concerns surrounding the sudden influx of individuals from diverse cultural backgrounds were perhaps propagated by disingenuous claims that resulted in the amplification of anti-refugee sentiment by specific actors. This potentially resulted in many primary concerns being redirected towards unwarranted fears and misconceptions concerning the perceived threat presented by the refugee population. The duration of the refugee's housing, particularly their announcement, is regarded as a trigger event. This chapter analyses the online activity related to the announcement, existence, and closure of the Penally asylum accommodation, focusing on general information flows and anti-refugee attitudes. Data collection spanned before during and after the refugees being housed in Penally between 1st August 2020 and 31st March 2021.

Online dynamics associated with trigger events pose novel challenges, especially regarding the impact of anonymity and the resulting deindividuation effects, which can significantly increase the likelihood of negative dissemination (Joinson, 1998). Investigating tensions during trigger events can provide new insights into how they manifest online after events and methodological contributions through comprehensive and granular approaches enabled by social media data. This

approach allows for a more reflective and insightful understanding not only of how anti-refugee attitudes manifest and are propagated but also of the precise timing and dynamics of these occurrences, ultimately improving the accuracy of reporting and enhancing our ability to track when and how such events unfold (Edwards et al., 2013). The chapter explores community tension patterns related to local issues over a gradual timescale, specifically focusing on the Penally asylum accommodation. It examines how various social actors, including communities, respond to sudden spikes in tensions and how they challenge and curtail further tensions. This analysis sheds light on the factors that enhance or further propagate tensions and information flows and the factors that contribute to constructive responses and the mitigation of them. To provide a comprehensive understanding of these patterns, the data is analysed using exploratory data analysis, statistical modelling, and qualitative extracts such as tweet string data, images, and aggregated word clouds. Additionally, Google perspective API (identity attack) is used to measure digital anti-refugee attitudes, providing initial evidence that this tool can be used in similar studies. Through this research, the chapter aims to contribute to the surrounding literature by revealing insights into anti-refugee attitudes and the factors that influence them and information dissemination more generally while also exploring ways to mitigate and challenge tensions.

## 5.2. Hypotheses

*H1: Offline events related to the Penally asylum accommodation will serve as a trigger event for increased online discussions, information flows and tensions.*

This study aims to investigate the prevalence of anti-refugee attitudes associated with the announcement, duration, and eventual closure of the Penally asylum accommodation. The research builds on previous studies that have shown an increase in tensions following antecedent trigger events, which galvanize and concentrate negative attitudes towards a particular social group. While most prior research has focused on such shocks in offline contexts (Legewie, 2013; King and Sutton, 2014; Hanes and Machin, 2014), this study extends knowledge of digital tensions drawing on recent research that has showed similar patterns in online

contexts (Ozalp et al., 2020; Czymara et al., 2022). The research question is exploratory, as the Penally asylum accommodation represent a trigger event that is less extreme and more localised than those previously examined in the literature. While research into trigger events has not been as extensively applied to local, culturally homogeneous areas, it is still hypothesised, based on Garland & Chakraborti's (2006) research, that these communities may not be immune to the emergence of negative discriminatory attitudes when faced with events disrupting cultural homogeneity.

*H2: Content, social, and external factors will be predictive of anti-refugee attitudes relating to the Penally asylum accommodation.*

The study aims to investigate the effect of six different content factors, namely mentions, hashtags, URLs, text sentiment, and text threat, on the production and extremality of anti-refugee tweets. Based on previous research, the study hypothesizes that URLs will be negatively associated with anti-refugee attitudes, whereas hashtags will be positively associated. This hypothesis is supported by research that suggests hashtags enable a process of "contagion" effect (Sunstein, 2017) for anti-refugee posts, while URLs are more likely to provide links to conventional news sources that are less likely to corroborate racist outlooks (Williams and Burnap, 2016)<sup>32</sup>. The study also hypothesis that positive sentiment will be negatively associated with anti-refugee attitudes, based on past findings (Ozalp et al., 2020). In other words, the more positive the tweet, the less likely it should involve negative attitudes targeted towards refugees. Finally, the study hypothesizes that threatening language will share a positive association with anti-refugee tweets. This hypothesis is based on previous research that examines language use during community tensions, which suggests that threatening language is, as perhaps expected more likely to be antagonistic (Roxell, 2011). Based on the findings of Williams and Burnap (2016), it is hypothesised that sentiment will be positively associated with information size flows, whereas anti-refugee attitudes will be

---

<sup>32</sup> Results should be interpreted cautiously, as they were obtained in 2016 and may be less applicable due to the rapidly evolving nature of digital environments.

negatively associated. No assumptions are made about the impact of threatening language.

The study also aims to examine the propagation of anti-refugee attitudes in relation to six agent types, including police, news, pro-refugee activists, anti-refugee activists, political, and far-right political. Based on previous research (Williams and Burnap; Ozalp et al., 2020), the study hypothesizes that all agent types will have a significant association with anti-refugee attitudes using the digital public as a reference category. Conventional media and political agents are hypothesised to be less associated with anti-refugee attitude propagation than the digital public (other agents), whereas anti-refugee activists and far-right politicians are hypothesised to be positively associated with the propagation of anti-refugee attitudes. The study also hypothesizes that the extent of the positive association with anti-refugee attitude propagation will be greater for anti-refugee activists, who predominately use aliases, due to deindividuation caused by perceived anonymity (Festinger, 1952). Two external factors, namely Google searches and news headlines, will be measured in relation to anti-refugee attitudes production. Cohen (2002: xxiii) posited that conventional channels of media are integral to agenda building in the case of deviant events through the use of images to set agendas and build narratives. This theory was further examined in digital contexts indicating a positive association between news headlines and the intensity of anti-refugee hatred, as well as the information dissemination that followed events (Williams & Burnap, 2016). However, no significant association was found between Google searches and anti-refugee attitudes in the previous study. Thus, the study hypothesises that Google searches will not significantly predict anti-refugee attitudes in the model.

*H3: Social media communication networks will be an integral component for anti-refugee content development.*

It is hypothesised dialogical factors on social media, such as mentions and replies will be integral tools for the production of anti-refugee attitudes. This will be explored in a number of ways, such as network visualisations of anti-refugee attitudes and information (size) propagation models by studying the 'mention' independent variable and dependent 'reply' variable. In terms of visualisations, it is hypothesised that

when filtered for just anti-refugee tweets, a core periphery structure will be evident (Borgatti and Everett, 2000), hence a small group of social actors will operate in a dense networked space, who account for the majority of anti-refugee content producers. This will be reflected in regression models; with dialogical metrics such as replies and mentions showing clear associations with anti-refugee attitudes. This hypothesis is based on surrounding literature that indicates the most negative and divisive attitudes manifest predominantly in networked 'echo-chambers' following trigger events (Müller & Schwarz, 2021).

*H4: Digital engagement metrics, such as likes, retweets, and replies, will associate with the degree to which anti-refugee attitudes and positive counter-speech are endorsed by specific social actors.*

Previous studies have linked information propagation and endorsement in digital spaces to on-platform metrics like retweets (Boyd et al., 2010), particularly in studies related to digital trigger events (Williams and Burnap, 2016; Ozalp et al., 2020). Collective efficacy, defined as "the process of activating or converting social ties among communities to achieve collective goals, such as public order or control of crime" (Sampson, 2010: 802), applies not only to various deviant behaviors but also to addressing community tensions (Petrosino and Pace, 2015). Sampson (2001) proposed that achieving collective efficacy often hinges on prominent information providers being perceived as unbiased and using non-inflammatory language. In the context of digital tensions, research has shown that conventional information providers like the police, news outlets, and MPs tend to have a greater likelihood of information propagation in terms of size and survival but are less likely to employ antagonistic language (Ozalp et al., 2020). Thus, it is hypothesised that information propagation will be more pronounced among political, news, and police agents. Community organisations actively challenging negative attitudes on social media have also been identified as significant contributors to collective efficacy, enjoying high endorsement by the digital public (Ozalp et al., 2020). Consequently, it is expected that anti-refugee activists will also experience high engagement and propagation, signifying their role in digital collective efficacy. However, prior research suggests that polarised actors motivated will exhibit a negative correlation with information propagation, indicating limited public endorsement. This includes anti-

refugee activists and far-right political figures (Williams and Burnap, 2016). No assumptions are made regarding polarised hashtags, whether pro or anti-refugee. Regarding external factors, no assumptions are made about the association between information propagation and Google searches. Nevertheless, Cohen's (2002) theory posits that the media continue to wield significant influence in setting agendas after significant events and shaping narratives. This theory aligns with contemporary studies showing that more modern media forms including online articles, maintain the ability to influence information flows remains relevant even in emerging digital contexts (Williams and Burnap, 2016).



## 5.3. Results

### 5.3.1. Descriptive Statistics<sup>33</sup>

Table 11. Descriptive Statistics (N= 6,661)

Variables	Coding	Sample	
		M/%	SD
<i>Dependent Variables</i>			
<i>Anti-refugee hate</i>	Range: 0.00-0.94	0.20	0.17
<i>Size (Retweets)</i>	Range: 0-562	2.98	17.78
<i>Size (Likes)</i>	Range: 0-2035	7.70	52.99
<i>Size (Replies)</i>	Range: 0-1628	1.71	25.03
<i>Independent Variables</i>			
<i>Content Factors</i>			
<i>Mentions</i>	Range: 0-49	1.04	1.78
<i>Hashtag</i>	Range: 0-16	0.40	1.13
<i>URL</i>	0 = no; 1 = yes	59%	-
<i>Word Count</i>	Range: 1-76	25.99	13.91
<i>Threat</i>	Range: 0.00-0.97	0.23	0.17
<i>Sentiment</i>	Range: 0-1	0.51	-
<i>Social Factors</i>			
<i>Pro-Hashtag</i>	0 = no; 1 = yes	2%	-
<i>Anti-Hashtag</i>	0 = no; 1 = yes	3%	-
<i>Ref: Non-Hashtag</i>	0 = no; 1 = yes	4%	-
<i>Police Agent</i>	0 = no; 1 = yes	>1%	-
<i>News Agent</i>	0 = no; 1 = yes	3%	-
<i>Pro-Refugee Activist</i>	0 = no; 1 = yes	1%	-
<i>Anti-Refugee Activist</i>	0 = no; 1 = yes	3%	-
<i>Political Agent</i>	0 = no; 1 = yes	1%	-
<i>Far-Right Political Agent</i>	0 = no; 1 = yes	2%	-
<i>Ref: Other Agent</i>	0 = no; 1 = yes	81%	-
<i>External Factors</i>			
<i>Google Searches</i>	Range: 0-100	17.36	23.94
<i>Press Headlines</i>	Range: 0-20	5.02	0.26
<i>Control</i>			
<i>August</i>	0 = no; 1 = yes	1%	-
<i>October</i>	0 = no; 1 = yes	24%	-
<i>November</i>	0 = no; 1 = yes	10%	-
<i>December</i>	0 = no; 1 = yes	8%	-
<i>January</i>	0 = no; 1 = yes	15%	-
<i>February</i>	0 = no; 1 = yes	7%	-
<i>March</i>	0 = no; 1 = yes	11%	-
<i>Ref: September (highest hate crime levels)</i>	0 = no; 1 = yes	25%	-

### **5.3.2. Exploratory Data Analysis (EDA)**

#### *5.3.2.1. Data Visualisation*

Prior to examining how content, social, external and control factors influence anti-refugee content and information diffusion relating to Penally, some of the data was visualised. This is useful in providing an exploratory account of how information flows and anti-refugee sentiments developed during the announcement of the asylum accommodation and how it differed from other trigger-events frequently studied in surrounding research. The findings were produced in dashboard format for practitioners, with seven visualisations (see Appendix 8). For the purpose of this thesis, they have been separated into separate figures.

Figure 15 provides an aggregated timeseries line graph for tweet frequency (blue line) and anti-refugee content (0.7+)<sup>34</sup> frequency (red line). The graph indicates that the tweet frequency related to the keyword 'Penally' began to increase on September 14th, when the housing of refugees in the Penally army barracks was announced. The tweet volume increased substantially, reaching its highest overall peak on September 22nd, nine days after the arrival and subsequent offline protests outside the barracks. The twelve-day peak period between September 14th and September 25th witnessed a significant rise in online 'Penally' tweet traffic, with a staggering 5692.12% increase in average tweets per day compared to the previous forty-five-day period. This sudden surge provides supporting evidence for the first hypothesis (H1) posited in previous research, which suggests that offline trigger events can lead to increased online information flows (Williams and Burnap, 2016; Ozalp et al., 2020; Czymara et al., 2022). Despite a visibly lower volume of anti-refugee content the red line shows a similar initial pattern following the overall tweet volume.

---

<sup>34</sup> The 0.7 threshold was determined by a combination of Perspective's guidelines (Google, 2021b) and human code showing a reliable F1 coefficient.

## A Case Study of the Penally Asylum accommodation

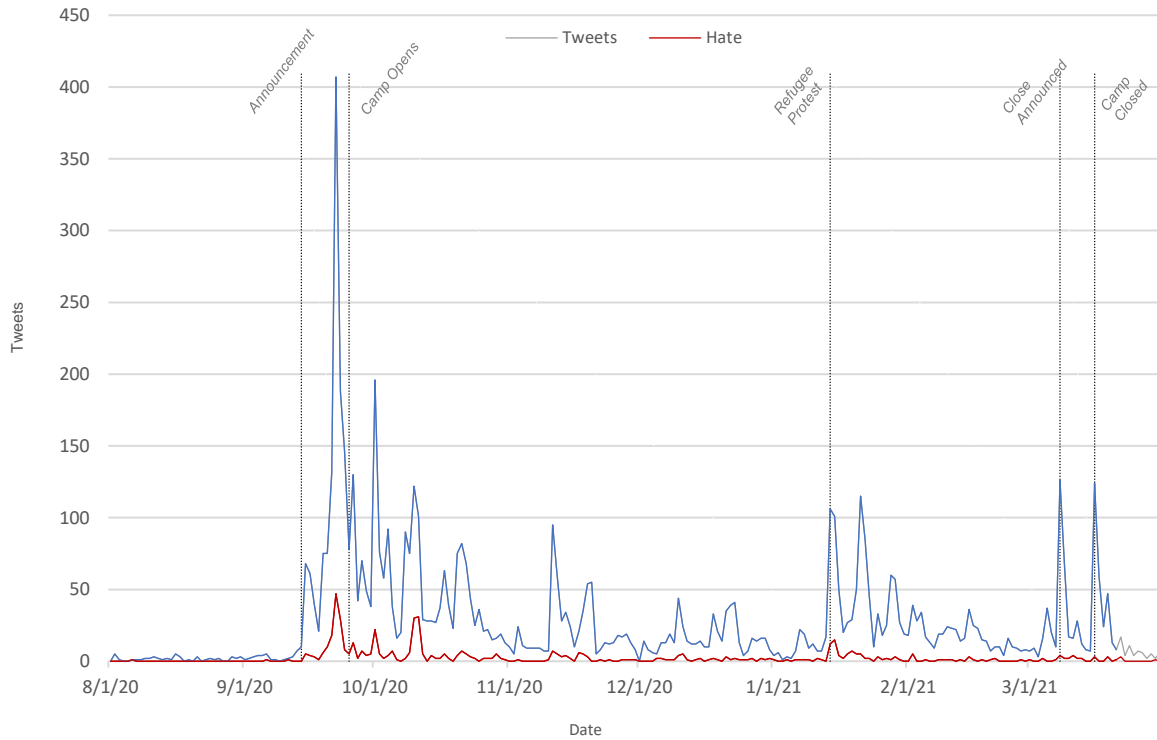


Figure 15: Time series for Penally Tweet traffic and anti-refugee posts

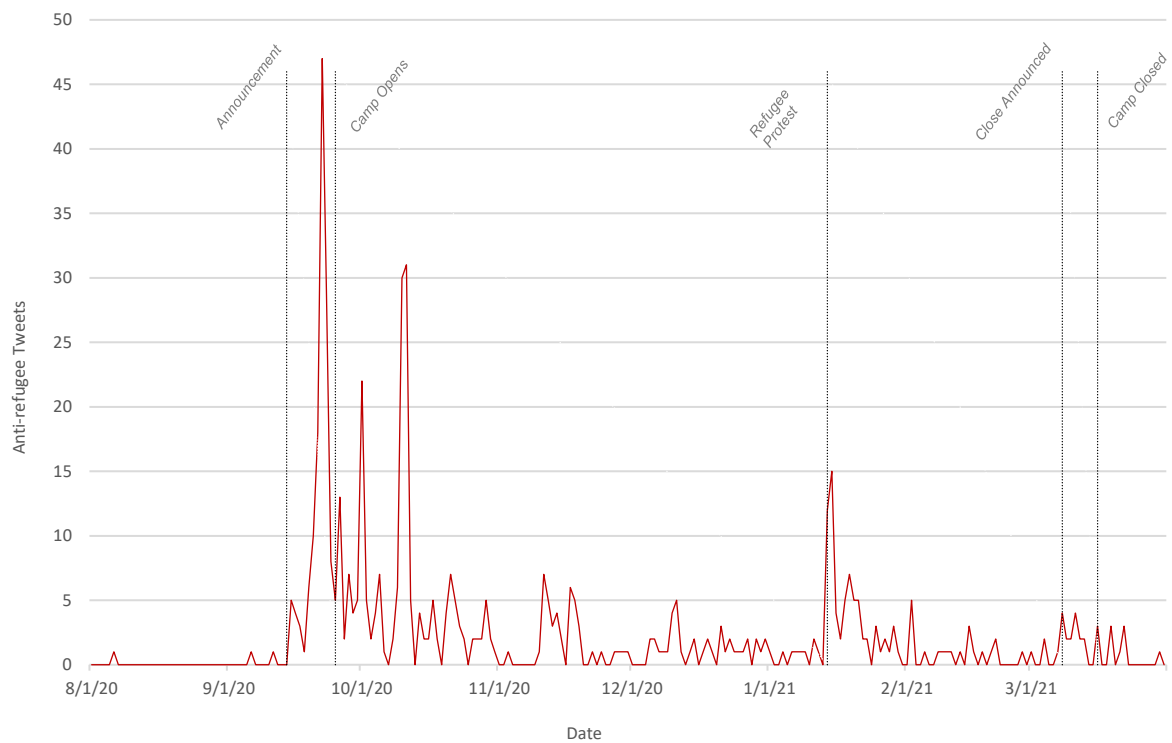
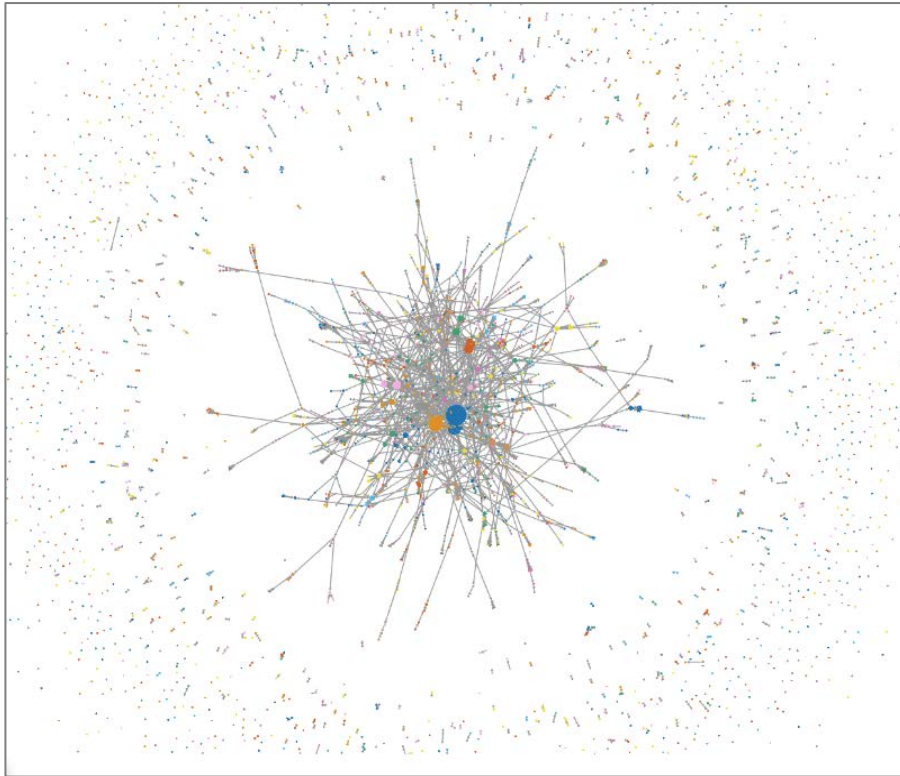


Figure 16: Time series for Penally anti-refugee posts

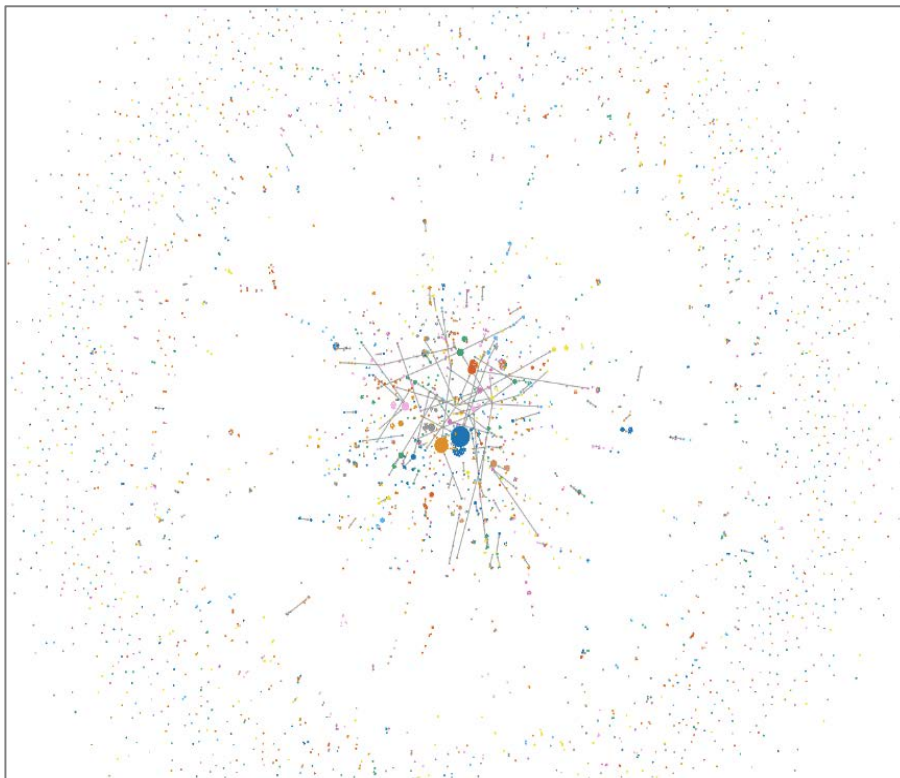
Figure 16 presents an enhanced and more comprehensible time series line graph depicting the prevalence of just anti-refugee tweets. Upon visual examination, discernible temporal correlations emerge between the overall tweet frequency (refer to Figure 15) and the occurrences of anti-refugee tweets. Notably, an initial surge is observed, wherein the twelve-day 'spike average' registers an 18,525% escalation compared to the preceding 45-day average. The zenith of anti-refugee content intensity is most pronounced immediately following the announcement, substantiating existing evidence that, despite the initial deleterious impact of heterogeneity on community cohesion, such effects tend to attenuate over time, with heightened acuity in the short term (Putman, 2007; Twigg et al., 2010).

Anti-refugee tweets consistently mirrored subsequent fluctuations, with a conspicuous spike occurring on January 14th during a public demonstration in the nearby town of Tenby, led by asylum seekers protesting against the living conditions in the camp. During this period, anti-refugee content exhibited synchronous escalation with general information flows, indicative of a resurgence phase (Chetty and Alathur, 2018), during which a notable segment of individuals resorted to online abuse directed at the refugees. This can be attributed, in part, to perceived threats to ingroup identity. The public demonstration by asylum seekers likely triggered heightened perceptions of threat among certain individuals, prompting them to express hostile sentiments online as a defensive response to perceived challenges to their ingroup cohesion.

Conversely, during the final surge in tweets in March, coinciding with the announcement and subsequent closure of the camp, instances of anti-refugee content did not parallel the overall trajectory of online discussions. This is evident in the discernible spikes in the blue line and the absence of corresponding spikes in the red line. This deviation may be elucidated by the proposition that individuals propagating anti-refugee sentiments may perceive such expressions as superfluous as the camps were being dismantled, albeit amid human rights concerns rather than as a consequence of anti-refugee protests.



*Figure 17: Reply node network for Penally (all replies)*



*Figure 18: Reply node network for Penally (anti-refugee tweet filter)*

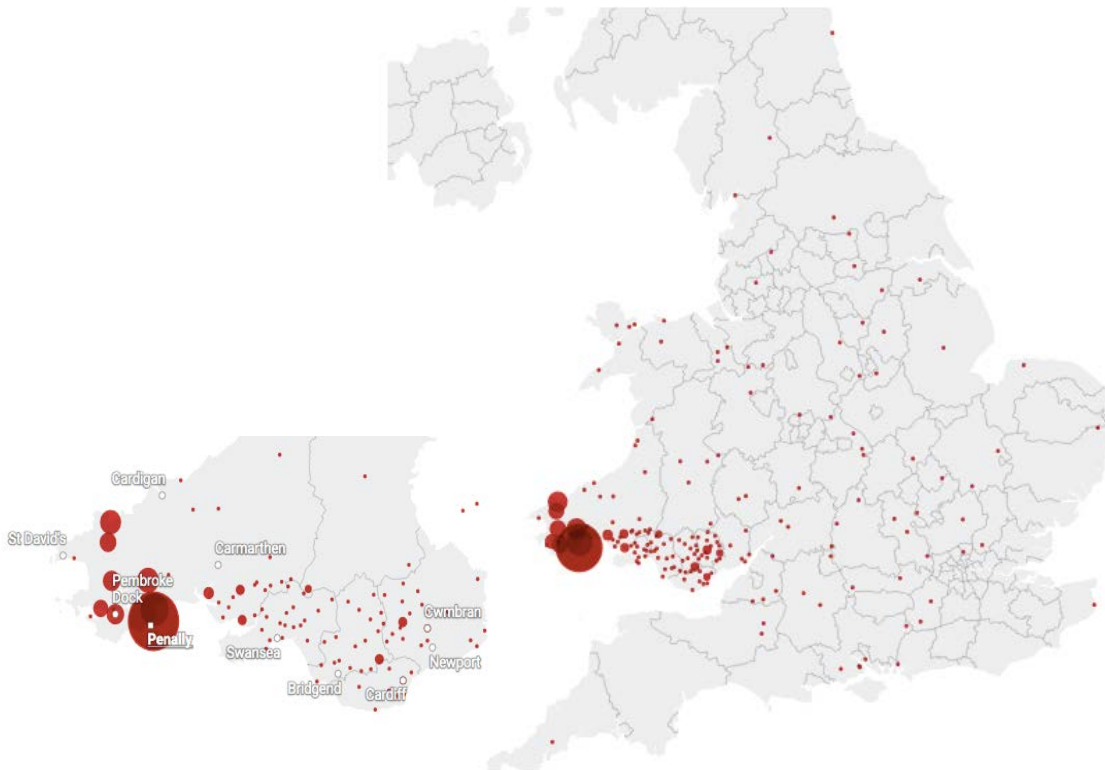
Figures 19 and 20 depict reply network visualisations from all tweets in the dataset ( $n=6,661$ ). Figure 17 illustrates all nodes in the dataset, with every reply being represented as an edge. While many tweets observed in the outer regions of the visualisation did not elicit any replies, a dense cooperation space existed in the centre, with many edges. This observation suggests that tweets pertaining to the Penally asylum accommodation displayed a core-periphery structure (Borgatti and Everett, 2000), which is a common phenomenon in specific focus issues in digital communities (Yang et al., 2018). While core-periphery structures are more prevalent in platforms such as Reddit, where separate pages are devoted to specific topics and events, this study provides evidence that such structures can exist on public space applications like Twitter following trigger events. Figure 18 likewise shows a census of nodes, but a filter was applied to better comprehend interaction patterns between propagators of anti-refugee tweets. Edges were shown only to reflect replies between tweets with an identity attack score above the study's threshold (0.7).

An initial inspection reveals that far fewer interaction patterns exist among anti-refugee tweets<sup>35</sup>. Interestingly, a core-periphery structure can still be identified among nodes with anti-refugee attitudes, indicating that a community of social actors in the dataset exists in a dense cooperation space who engage with anti-refugee attitudes relating to refugees in Penally. This finding offers corroborative evidence in line with previous studies, indicating the possibility of the emergence of "echo-chambers" (Müller and Schwarz, 2021), whereby like-minded individuals engage in interactions within a dense core (Ozalp et al., 2020). In this case, a small minority of social actors operate within a tightly knit network, which is responsible for the majority of anti-refugee content witnessed during the study period. The visualisations are further supported by bivariate analysis, which indicates that 2.41% of accounts produced all of the most extreme anti-refugee tweets (above 0.7) during the course of the study. This outcome provides initial support for the hypothesis that discussion networks are a crucial tool for the development and propagation of anti-refugee

---

<sup>35</sup> This outcome is expected, given the significantly lower volumes of identity attack tweets depicted in the time-series graphs.

attitudes (H6). Later in the study regression models are used to further explore the relationship between dialogical tools such as mentions and anti-refugee attitudes.



*Figure 19: Google search trend heat map for "Penally Camps" keyword*

Figure 19 displays heat maps of online information flows concerning Penally. Previous studies have utilised API geo-tagging features on platforms like COSMOS to comprehend the geographic basis of on-platform information flows (Williams and Burnap, 2016). However, this approach is limited since only 1% of tweets are geo-tagged due to user preferences (Sloan et al., 2013). In order to address this limitation, this study instead used Google search trend data to geo-tag online information flows. While this approach does not directly relate to active discourse flows on social media, it provides a more reflective overview of online interest in Penally during the study period. The UK heatmap (right side) indicates that online interest is mainly concentrated in Wales<sup>36</sup>, and specifically south Wales. While some

---

<sup>36</sup> Larger markers denote higher accumulative searches, whereas smaller markers indicate lower accumulative searches.





later analysis. Words such as 'community' and 'local' further support previous findings from the geo-tag visualisations, emphasising Penally's more localised nature compared to other trigger events. Later word cloud aggregations are used to explore specific social actors and themes, providing a comprehensive understanding of discrepancies in the data. Overall, the visualisations provide an opportunity to position the Penally case study in relation to other trigger event examples. Online responses to the Penally asylum accommodation differ notably from those observed in other trigger events. Unlike the typically short-lived online activity following events like terrorist incidents, homicides, and political votes, which generally last for a few days to a month, the online activity related to the Penally accommodation persists over an extended period. This extended duration aligns with the prolonged nature of the event itself. This results in identifiable 'rebirth stages' that are not assumed in other trigger events. This study fills a significant gap in the literature by exploring how online activity and antagonistic content manifests in the wake of a non-extremist, localised and longer-spanning trigger event. Overall, the visualisations provide valuable insights into public sentiment and key agents in the tension surrounding the Penally asylum accommodation, while also highlighting the unique nature of this trigger event compared to others.

#### *5.3.2.2. Agent Comparison*

Prior to conducting regression analyses, a Kruskal Wallis H test was employed to compare differences between social agents. Table 12 displays the mean rank statistics for each agent type for the three content factors, providing initial insights into the discrepancies in content factors used. This analysis sheds light on how social actors employ social media communications to disseminate information following the Penally trigger event. The initial rankings reveal limited discrepancies between agents in their use of hashtags, except for pro-refugee activists who used hashtags more frequently than any other agent. Post hoc comparisons tests for agent type indicated statistically significant differences between pro-refugee activists

and all other agent types<sup>38</sup>. Regarding mentions, far-right political and anti-refugee activists ranked significantly higher in usage than the other agents. Further post hoc tests between anti-refugee activists<sup>39</sup> and far-right political agents<sup>40</sup> indicated statistically significant differences in mention use when compared to all other agent types except the police ( $p=1.02/ p= 0.98$ ). However, no statistically significant differences were observed in mention use between far-right political and anti-refugee activists ( $p= 1.12$ ). For hyperlinks, more established and traditional institutions such as the police, news, and political agents outranked remaining social actors.

*Table 12. Agent comparison of modalities*

<i>Agent Type</i>	<i>Hashtag</i>	<i>Mention</i>	<i>Hyperlink</i>
<i>Far Right Political</i>	3522.28	4772.71	2407.39
<i>Police</i>	3166.2	3995.88	4694
<i>News</i>	3461.41	1859.26	4596.04
<i>Anti-Refugee Activist</i>	3449.58	4784.07	3213.78
<i>Pro-Refugee Activist</i>	4333.53	3960.1	3373.28
<i>Political</i>	2769.75	2072.25	4542.61
<i>Digital Public</i>	3314.68	3321.18	3306.7
<i>Sig</i>	0.00	0.00	0.00
<i>Kruskal-Wallis H</i>	48.97	264.79	180.78

<sup>38</sup> Far right political ( $p= <.01$ ), police ( $p= <.01$ ), news ( $p= <.01$ ), anti-refugee activists, political ( $p= <.01$ ), and digital public ( $p= <.01$ ).

<sup>39</sup> Police ( $p= <.01$ ), news ( $p= <.01$ ), political ( $p= <.01$ ), and digital public ( $p= <.01$ ).

<sup>40</sup> Police ( $p= <.01$ ), news ( $p= <.01$ ), political ( $p= <.01$ ), and digital public ( $p= <.01$ ).

5.3.2.3. Anti-refugee attitudes model

Table 13. OLS regression predicting anti-refugee attitudes (Penally)

	<i>B</i>	<i>SE</i>	<i>b</i>
<i>Content Factors</i>			
<i>Mentions</i>	0.00	0.00	0.05***
<i>Hashtag</i>	0.01	0.00	0.05***
<i>URL</i>	-0.02	0.00	-0.06***
<i>Word Count</i>	0.00	0.00	0.13***
<i>Threat</i>	0.47	0.01	0.45***
<i>Sentiment</i>	-0.08	0.01	-0.12***
<i>Social Factors</i>			
<i>Anti-Hashtag</i>	0.11	0.01	0.09***
<i>Pro-Hashtag</i>	0.05	0.02	0.04***
<i>Ref: Non-Hashtag</i>			
<i>Police Agent</i>	-0.09	0.04	-0.02**
<i>News Agent</i>	-0.03	0.01	-0.03***
<i>Pro-Refugee Activist</i>	0.01	0.02	0.01
<i>Anti-Refugee Activist</i>	0.13	0.01	0.11***
<i>Political Agent</i>	-0.11	0.02	-0.05***
<i>Far-Right Political Agent</i>	0.03	0.01	0.03***
<i>Ref: Other Agent</i>			
<i>External Factors</i>			
<i>Google Searches</i>	0.00	0.00	0.02
<i>Press Headlines</i>	0.00	0.00	0.02
<i>Control</i>			
<i>August</i>	-0.08	0.02	-0.08***
<i>October</i>	-0.01	0.01	-0.04**
<i>November</i>	-0.03	0.01	-0.04***
<i>December</i>	0.00	0.01	-0.01*
<i>January</i>	-0.03	0.01	-0.05***
<i>February</i>	-0.04	0.01	-0.05***
<i>March</i>	-0.05	0.01	-0.09***
<i>Ref: September (highest hate crime levels)</i>			
<i>Model Fit</i>			
<i>Sig</i>			0.00
<i>R2</i>			.31
<i>n</i>			6,661

Table 13 shows the results of the anti-refugee content regression model<sup>41</sup>. All five content predictors returned significant results. Firstly, hyperlinks were found to have a negative association with anti-refugee attitudes, while hashtags were found to have a positive association. These results are consistent with previous research, which has suggested that individuals seeking to propagate antagonistic content online may utilise hashtags to increase their content's visibility, potentially leading to a contagion effect (Sunstein, 2017).

Tweets that utilised threatening language (as measured by Perspective API) were positively associated with anti-refugee attitudes. This result is not surprising, as threatening language is often used in tandem with discriminatory content (Roxell, 2011). Tweets with more mentions were positively associated with anti-refugee attitudes. This finding supports earlier visualisations that indicated that dialogical processes play a significant role in forming networked dense echo chambers where like-minded people can share and endorse each other's anti-refugee opinions. Sentiment was the only other content factor that was negatively associated with anti-refugee tweets, indicating that tweets with more positive language were less likely to involve anti-refugee attitudes. This outcome is somewhat anticipated, as positive language is unlikely to be antagonistic<sup>42</sup>. Prevalent hashtags were categorised into pro-refugee and anti-refugee, with "non-hashtags" serving as the reference category. The non-hashtags category included tweets with hashtags unrelated to the refugee issue, as well as tweets without any hashtags. Surprisingly, both pro-refugee and anti-refugee hashtags were positively associated with identity attack content in tweets. Reasons for this are explored in the qualitative section later in this chapter.

Agent type was explored, using the digital public ("non-agent") as the reference category, which constituted 91.1% of the dataset. This cohort reflects the concept of

---

<sup>41</sup> The dataset used in this analysis represents a near-census of posts related to the subject under investigation. As a result, it is not surprising that a significant majority of coefficients exhibit high levels of statistical significance. Therefore, the primary emphasis of this chapter's analysis is on elucidating the magnitude and direction of associations rather than dwelling solely on the statistical significance of individual coefficients. (applicable to remaining models in this chapter)

<sup>42</sup> This finding corroborates the reliability and validity of the identity attack perspective API classifier.

a digital public agora (Burnap et al., 2015). Significant associations were found between identity attack tweets and all alternative agents except for pro-refugee activists. The coefficients for the remaining social agents are presented in Figure 21.

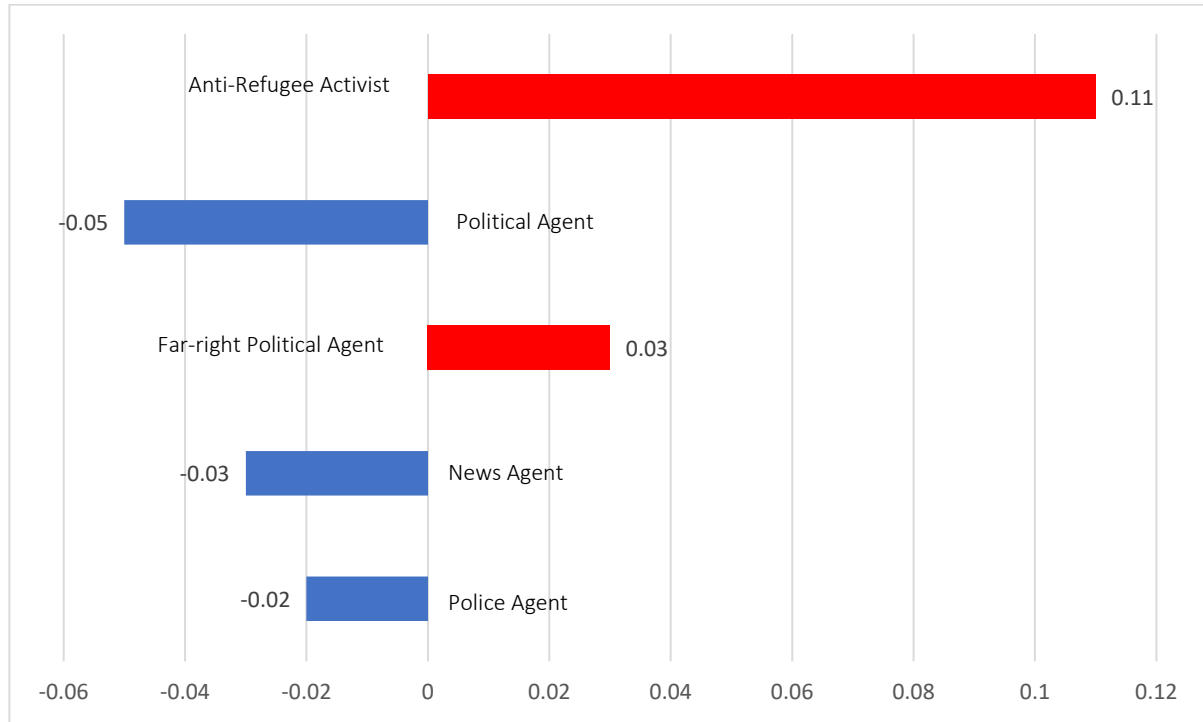


Figure 21. Anti-refugee attitudes coefficients by agent<sup>43</sup>

A quick inspection of Figure 21 shows anti-refugee activists and far-right political agents are the only two agent-types to have positive associations with anti-refugee attitudes. Once again, this reaffirms the effectiveness of the perspective API classifier in accurately identifying antagonistic speech. Conversely, political, police and news agents were found to have negative associations with anti-refugee attitudes in relation to the digital public. This plot provides a better understanding of the extent to which different agent types associate with anti-refugee attitudes. Specifically, anti-refugee activists were significantly more likely to author anti-refugee tweets than any other agent, with far-right political agents being more associated with anti-refugee attitudes than the general digital public, although to a lesser extent

<sup>43</sup> Reference category is “Other agent” (digital public). Red bar = positive association with anti-refugee attitudes, blue bar = negative association with anti-refugee attitudes. All coefficients significant at the P Value (<0.05)

than anti-refugee activists. The negative coefficients for the remaining agents indicate that the digital public are more likely to produce anti-refugee content than any of the traditional institutions. Neither of the two external factor predictors (google searches and press headlines) had significant associations with anti-refugee attitudes, lending further evidence to established agencies being less likely to contribute to anti-refugee attitudes following trigger events.

All months, except for December, produced significant negative values. Supporting H2, all other months were negatively associated with identity attacks in relation to the reference month September. This suggests that, when controlling for all other factors, the month with the highest levels of offline hate also saw the highest levels of anti-refugee content online. This notion is further supported when examining the coefficients for the remaining months in relation to the reference month September. Specifically, the coefficients for October and November, which had high amounts of police-recorded offline hate had the joint weakest negative associations with anti-refugee attitudes. Therefore, while October and November were still less likely to predict high levels of anti-refugee attitudes than September, the extent of this difference was much less than other months (with lower levels of offline hate).

#### *5.3.2.4. Sub-Factor Analysis*

In addition to examining individual predictors, a sub-factor analysis was conducted to determine which set of variables accounted for the most variance in the model. The adjusted R2 metric was employed to determine the variance caused by the four sub-factors. External factors, which did not produce any significant predictors, accounted for the least variance in the model (R2 .008). The second least amount of variance was explained by controls, namely the months in which tweets were authored (R2 .023). Social factors accounted for the second most variance in the model (R2 .048), while content factors explained the most anti-refugee attitude variance by a significant margin (R2 .279). The results of the sub-factor analysis suggest that although all factor groups can impact levels of anti-refugee attitudes, it is primarily influenced by the tweet's content, with a relatively large amount also attributed to the agent who authors it.

#### 5.3.2.5. *Digital Engagement Models*

The findings from each of the negative binomial models (refer to Table 14, 16, and 17) reveal numerous statistically significant links between the dependent variable (digital engagement metrics) and the predictive factors. These connections exhibit varying strengths, as the incidence rate ratio (IRR) indicates. Given the high prevalence of statistically significant results, potentially influenced by the dataset's near-census coverage of tweets, the analysis primarily focuses on assessing the magnitude and directions of these relationships by considering the IRR value<sup>44</sup>.

---

<sup>44</sup> The IRR is calculated by exponentiating the coefficients obtained from the negative binomial regression. This allows for interpretation of engagement incidence rates, rather than dealing with logs of expected engagement counts. Consequently, the IRR serves as a valuable tool for identifying associations between specific factors and digital engagement, allowing for an identification of which factors are more relevant than others.

Table 14: Negative binomial regression, predicting like counts (size model)

	<i>B</i>	<i>SE</i>	<i>Wald</i>	<i>IRR</i>
<i>Content Factors</i>				
<i>Mentions</i>	-0.14	0.01	364.39	0.87***
<i>Hashtag</i>	-0.11	0.01	66.93	0.89***
<i>URL</i>	-0.17	0.03	30.17	0.85***
<i>Word Count</i>	0.05	0.00	2109.20	1.05***
<i>Threat</i>	-0.30	0.10	8.75	0.74***
<i>Sentiment</i>	0.56	0.05	120.43	1.76***
<i>Anti-refugee hate</i>	0.98	0.10	102.74	2.65***
<i>Social Factors</i>				
<i>Pro-Hashtag</i>	-0.12	0.30	0.15	0.89
<i>Anti-Hashtag</i>	-0.20	0.29	0.47	0.82
<i>Ref: Non-Hashtag</i>				
<i>Police Agent</i>	0.32	0.32	1.02	1.38
<i>News Agent</i>	-0.66	0.08	61.96	0.52***
<i>Pro-Refugee Activist</i>	0.67	0.16	18.28	1.96***
<i>Anti-Refugee Activist</i>	0.09	0.09	0.92	1.09*
<i>Political Agent</i>	1.68	0.18	84.66	5.36***
<i>Far-Right Political Agent</i>	-1.61	0.09	304.21	0.20***
<i>Ref: Other Agent</i>				
<i>External Factors</i>				
<i>Google Searches</i>	0.00	0.00	0.43	1.00
<i>Press Headlines</i>	0.04	0.00	93.71	1.04***
<i>Control</i>				
<i>August</i>	-0.31	0.17	3.20	0.73*
<i>October</i>	0.11	0.05	5.83	1.12**
<i>November</i>	-0.27	0.06	18.68	0.76***
<i>December</i>	0.35	0.07	24.57	1.41***
<i>January</i>	0.18	0.06	10.86	1.20***
<i>February</i>	-0.22	0.07	10.45	0.80***
<i>March</i>	0.04	0.06	0.40	1.04
<i>Ref: September (highest hate crime levels)</i>				

Table 14 presents the outcomes of the like engagement model, which aimed to examine the relationship between various factors and the number of likes received by tweets. The results revealed that among the content factors, the use of mentions (IRR= 0.87), hashtags (IRR= 0.89), and hyperlinks (IRR= 0.87) had similar negative association with likes, while longer tweets (IRR= 1.05) exhibited a slight positive association. Interestingly, tweets with threatening language (IRR= 0.74), had a negative association with the like count, while both positive sentiment (IRR= 1.76) and to a greater extent anti-refugee attitudes (IRR= 2.65), had a positive association. This finding suggests that tweets with polarising sentiment, both negative and positive, regarding refugees are more likely to receive likes, indicating a highly



polarised environment. Concerning social factors, prevalent hashtags related to pro-refugee and anti-refugee sentiments did not demonstrate significant associations with like count. However, except for police, most agent types returned significant associations. Political agents (IRR= 5.36) showed an extremely strong positive association with likes, while both pro-refugee (IRR= 1.96), and to a slightly lesser extent anti-refugee (IRR= 1.09), activists also had positive associations. These findings imply that polarised tweets authored by both pro and anti-activists<sup>45</sup> are associated with likes. Far-right political agents showed an extremely negative association with likes (IRR= 0.20). In terms of external factors, google searches did not show any significant association with likes, while press headlines had a slightly positive association. Finally, considering the months in which tweets were authored, tweets posted in August, November, and February were less likely to receive likes compared to those posted in September (the reference month with most anti-refugee attitudes), whereas those posted in October, December, and January were more likely to receive likes. The significant temporal associators with engagement support surrounding work<sup>46</sup> (Burnap et al., 2014; Zarrella, 2009).

---

<sup>45</sup> Anti-refugee activist predictor only approached conventional levels of statistical significance.

<sup>46</sup> This finding is consistent across all three engagement models.

Table 15. Negative binomial regression, predicting retweet counts (size model)

	<i>B</i>	<i>SE</i>	<i>Wald</i>	<i>IRR</i>
<i>Content Factors</i>				
<i>Mentions</i>	-0.13	0.01	228.10	0.88***
<i>Hashtag</i>	-0.07	0.02	19.50	0.93***
<i>URL</i>	-1.11	0.04	937.38	0.33***
<i>Word Count</i>	0.06	0.00	1820.17	1.06***
<i>Threat</i>	-0.39	0.12	10.86	0.67***
<i>Sentiment</i>	0.41	0.06	41.95	1.51***
<i>Anti-refugee hate</i>	-0.17	0.011	90.49	0.85***
<i>Social Factors</i>				
<i>Pro-Hashtag</i>	-0.15	0.32	0.23	0.86
<i>Anti-Hashtag</i>	-0.40	0.31	1.63	0.67
<i>Ref: Non-Hashtag</i>				
<i>Police Agent</i>	0.42	0.35	1.44	1.52
<i>News Agent</i>	-0.74	0.09	71.22	0.48***
<i>Pro-Refugee Activist</i>	0.52	0.18	8.67	1.68***
<i>Anti-Refugee Activist</i>	-0.55	0.10	31.59	0.58***
<i>Political Agent</i>	1.00	0.19	28.61	2.71***
<i>Far-Right Political Agent</i>	-1.86	0.10	357.05	0.16***
<i>Ref: Other Agent</i>				
<i>External Factors</i>				
<i>Google Searches</i>	0.00	0.00	0.19	1.00
<i>Press Headlines</i>	0.04	0.00	60.67	1.04***
<i>Control</i>				
<i>August</i>	1.69	0.41	17.01	5.43***
<i>October</i>	-0.18	0.05	10.88	0.84***
<i>November</i>	-0.18	0.07	6.45	0.83***
<i>December</i>	0.25	0.08	9.63	1.28***
<i>January</i>	-0.08	0.06	1.61	0.92
<i>February</i>	-0.26	0.08	11.65	0.77***
<i>March</i>	-0.11	0.07	2.29	0.90
<i>Ref: September (highest hate crime levels)</i>				

Table 15 presents findings from the size model for retweets. Content factors such as mentions (IRR= 0.88), hashtags (IRR= 0.93), and to a greater extent URLs (IRR= 0.33) were negatively associated with retweets, while longer tweets (IRR= 1.06) were positively associated. Additionally, tweets using threatening language (IRR= 0.67) were also negatively associated with retweets. Interestingly, anti-refugee attitudes (IRR= 0.85) had a negative association with retweets, in contrast to its positive association with likes. Similar to likes, sentiment (IRR= 1.51) was positively associated with retweets. Prevalent hashtag types (pro and anti-refugee) did not significantly affect retweet count. Only police agents returned non-significant associations with retweet count among agent types. Far-right political agents (IRR=

0.16) exhibited the greatest negative association with retweet count followed by news agents (IRR= 0.48) and anti-refugee activists (IRR= 0.58). Conversely, pro-refugee activists (IRR= 1.68) and to a greater extent political agents (IRR= 2.71) were positively associated. Regarding external factors, google searches were not significantly associated with retweets, while news headlines were minimally positively associated. Finally, tweets authored in August and December were more likely to be retweets than those posted in September (the reference category). In contrast, those authored in October, November, and February were less likely to be retweets.

Table 16. Negative binomial regression, predicting reply counts (size model)

	<i>B</i>	<i>SE</i>	<i>Wald</i>	<i>IRR</i>
<i>Content Factors</i>				
<i>Mentions</i>	-0.20	0.01	216.07	0.82***
<i>Hashtag</i>	-0.11	0.02	36.27	0.89***
<i>URL</i>	-0.84	0.04	381.36	0.43***
<i>Word Count</i>	0.05	0.00	1075.76	1.05***
<i>Threat</i>	-0.83	0.13	38.79	0.44***
<i>Sentiment</i>	-0.24	0.07	10.93	0.79***
<i>Anti-refugee hate</i>	1.87	0.12	263.06	6.52***
<i>Social Factors</i>				
<i>Pro-Hashtag</i>	1.38	0.56	6.05	3.96**
<i>Anti-Hashtag</i>	-1.08	0.56	3.77	0.34**
<i>Ref: Non-Hashtag</i>				
<i>Police Agent</i>	1.55	0.54	8.38	4.73***
<i>News Agent</i>	-0.75	0.10	59.88	0.47***
<i>Pro-Refugee Activist</i>	0.64	0.24	6.99	1.90***
<i>Anti-Refugee Activist</i>	0.60	0.14	19.17	1.82***
<i>Political Agent</i>	2.24	0.32	48.89	9.43***
<i>Far-Right Political Agent</i>	-1.66	0.11	225.07	0.19***
<i>Ref: Other Agent</i>				
<i>External Factors</i>				
<i>Google Searches</i>	0.00	0.00	8.00	1.00***
<i>Press Headlines</i>	0.06	0.01	108.80	1.06***
<i>Control</i>				
<i>August</i>	-0.21	0.27	0.61	0.81
<i>October</i>	0.06	0.06	0.96	1.06
<i>November</i>	-0.79	0.08	106.29	0.45***
<i>December</i>	0.67	0.10	43.92	1.95***
<i>January</i>	0.09	0.07	1.59	1.10
<i>February</i>	-0.69	0.08	67.80	0.50***
<i>March</i>	0.61	0.09	48.64	1.83***
<i>Ref: September (highest hate crime levels)</i>				

Table 16 presents results from the size model for replies, which investigated the association between various factors and the number of replies received by tweets. Among content factors, mentions (IRR= 0.82), hashtags (IRR= 0.89), and to a greater extent hyperlinks (IRR= 0.43) had negative associations with replies, while word count (IRR= 1.05) had a slight positive association. Interestingly, tweets with threatening (IRR= 0.44) or to a lesser extent positive sentiment (IRR= 0.79) were negatively associated with replies. In contrast, anti-refugee attitudes (IRR= 6.53) was positively associated with a large magnitude, suggesting that such tweets may provoke more discussion. All social factors exhibited significant associations with replies. Pro-refugee hashtags had an extremely positive association with replies (IRR= 3.96). In contrast, anti-refugee hashtags (IRR= 0.34) had a negative association, indicating that hashtags were more effective for propagating discussion around pro-refugee information than anti-refugee information. In terms of agents, political agents (IRR= 9.43) had an extremely positive association with replies. In contrast, the police (IRR= 4.73), pro-refugee activists (IRR= 1.90), and anti-refugee activist (IRR= 1.82) also had positive associations (although to a lesser extent). This suggests that politicians were effective in instigating online discussion about the Penally asylum accommodation, while far-right political agents and news agents had negative associations when compared to the digital public. Regarding external factors, an increase in press headlines (IRR= 1.06) was positively associated with increased replies, although to a minimal extent. Finally, when considering the months in which tweets were authored, tweets posted in December and March were more likely to receive replies than those published in September (the reference month), whereas those posted in November and February were less likely to receive replies. No significant associations were found for tweets posted in August, October, and January.

### **5.3.3. Sub-model Analyses**

Sub-model analyses were carried out between the factor sets for each of the three engagement models (likes, retweets, and replies). Much like the sub-model analysis for the anti-refugee content model, the factor sets were manually determined rather than inferred from PCA factor reduction techniques. Appendix 9 indicate how much

variance is determined by each set for each model. For likes<sup>47</sup>, content factors explained the most variance, followed by social factors, **external factors** and finally, controls. For retweets<sup>48</sup>, content factors explained the most variance, followed by social factors, **external factors** and finally controls. For the size model relating to replies,<sup>49</sup> content factors explained the most variance, followed by controls, social factors, **and finally external factors**.

In summary, the sub-factor analyses of the three user engagement models revealed that content factors (mentions, hashtags, URLs, wordcount, threat, sentiment, and identity attack) accounted for the most significant variance. Social factors (prevalent hashtags and agent type) explained the second most variance in the like and retweet models but was third in the reply model. External factors (google searches and press headlines) explained the third amount of variance in both the like and retweet models, but the least in the reply model. Control factors (month tweet was published) explained the least amount of variance for the retweet and like models, but the second most for the reply model. Notably, the sub-factor findings for the like and retweet models were identical, suggesting a similar effect size of the selected regression factors on these metrics. However, the sub-model analysis for replies revealed that they interact differently with the variables in this study compared to other user engagement metrics.

---

<sup>47</sup> **Likes Model:** content factors explained the most variance (-2 Log-Likelihood = **-18787.98**, BIC= **37646.34**), followed by social factors (-2 Log-Likelihood = **-20085.00**, BIC= **40258.04**), **external factors** (-2 Log-Likelihood = **-20479.56**, BIC=**40985.54**) and finally controls (-2 Log-Likelihood = **-20515.15**, BIC= **41100.72**).

<sup>48</sup> **Retweets Model:** content factors explained the most variance (-2 Log-Likelihood = **-12888.53**, BIC= **25847.44**), followed by social factors (-2 Log-Likelihood = **-14229.68**, BIC= **28547.40**), **external factors** (-2 Log-Likelihood = **-14810.59**, BIC=**29647.59**) and finally controls (-2 Log-Likelihood = **-14838.31**, BIC= **29755.06**).

<sup>49</sup> **Replies Model:** content factors explained the most variance (-2 Log-Likelihood = **-10215.39**, BIC= **20501.15**), followed by controls (-2 Log-Likelihood = **-11386.93**, BIC= **22844.29**), social factors (-2 Log-Likelihood = **-11592.59**, BIC= **23273.23**) **and finally external factors** (-2 Log-Likelihood = **-11724.47**, BIC=**23475.35**).

## 5.4. Discussion

### 5.4.1. *The relationship between online and offline tensions*

Findings in this chapter provide support for H1, which suggests that offline events, such as the announcement of the Penally asylum accommodation, can trigger a significant increase in both general discourse and anti-refugee attitudes in digital spaces. Specifically, during the 12-day spike period following the announcement, there was a 5692.12% increase in general online discourse and a 18525% increase in anti-refugee attitudes compared to the previous 45 days. This study supports Garland & Chakraborti's (2006) findings by demonstrating that negative discriminatory attitudes can emerge in culturally homogeneous rural communities and further this knowledge by exploring them in the aftermath of antecedent trigger events.

Time-series visualisations further corroborated these findings, revealing additional spikes in online discourse and anti-refugee attitudes related to offline protests and moments of significance, such as the announcement of the camp's closure. This finding is significant because it highlights the potential for such sentiments to undergo a rebirth stage (Chetty and Alathur, 2018) that can occur multiple times and last longer in cases of localised policy issues compared to more extreme nationwide triggers such as one-time elections or extremist attacks. Notwithstanding, the timeseries visualisations revealed that the intensity of anti-refugee content was highest in the immediate aftermath of the announcement. This finding supports the evidence that, despite the initial negative impact of heterogeneity on community cohesion, the effects can diminish over time and are most acute in the short-term. This is consistent with the observations made by Putman (2007) and Twigg et al. (2010).

Statistical measures were employed to further explore the relationship between offline incidents and anti-refugee content. The study included monthly offline police-recorded hate crime and incident data from the local police region (Dyfed-Powys) in the anti-refugee attitudes (identity attack) regression model, with September being selected as the reference category due to its highest police-reported levels of hate

crimes and incidents. The results lend further support to H1, as all other months were less likely to generate anti-refugee attitudes when compared to September, which was also the anti-refugee attitudes peak in the time-series graph (the twelve-day spike). Moreover, this study builds upon previous research suggesting that offline trigger events can kickstart and sustain spikes in digital tensions, even in less extreme and more localised cases (Hanes and Machin, 2014; King and Sutton, 2014; Williams and Burnap, 2016). Consistent with previous studies examining digital tension trigger events (Williams and Burnap, 2016-1% and Ozalp et al., 2020-0.7%), severe anti-refugee attitudes on Twitter only accounted for 1.5% of all tweets following the event, despite clear instances of anti-refugee content being evident online.

#### *5.4.1.1. Understanding anti and pro-refugee posts*

After confirming the presence of anti-refugee attitudes in relation to camps, the analysis aimed to uncover the underlying factors associated with its use. Supporting H2, content and social factors were significant predictors of anti-refugee attitudes. Through sub-model analyses, preliminary evidence was obtained regarding the most causative factor groups of anti-refugee attitudes, ranked in order of importance. The models revealed that content factors accounted for the highest variance, followed by social, controls and external factors, indicating that factors related to the content of the tweets played a crucial role in developing anti-refugee attitudes. These findings align with previous research examining trigger events in digital spaces, such as Ozalp et al. (2020) studying antisemitic tensions. The nature of individual content factor predictors also corroborated earlier work, with hashtags having a positive association with anti-refugee attitudes, while hyperlinks had a negative association. This latter finding is consistent with similar studies on digital trigger events, as Williams and Burnap (2016: 227) suggest that hyperlinks may be less common in antagonistic tweets since "linked content (most often a popular media source) is unlikely to corroborate racist opinion and biased speculative rumours." A brief examination of some of the hyperlinks supported this notion, as many provided links to traditional institutions that were less predictive of anti-refugee attitudes in the models, such as news outlets like BBC articles or police press statements relating to

the camps. Consistent with recent research, antagonistic content was unsurprisingly found to have a negative correlation with positive sentiment (Kyrollos & Green, 2021).



Figure 22. Positive sentiment tweets (0.7+) (top 40 words, with numbers, special characters and stop words removed)

Upon closer examination of tweets that exhibited high levels of positive sentiment, it was determined that they were non-trivial and conveyed a pro-refugee sentiment<sup>50</sup>. Therefore, positive sentiment can serve as an indicator of the pro-refugee constituency within the digital public. No assumptions were made about the relationship between anti-refugee attitudes and tweet length. Still, results revealed that longer tweets tended to contain more anti-refugee attitudes, implying that social actors who promote anti-refugee attitudes tend to use longer format tweets.

A positive association was found between the use of threatening language and anti-refugee attitudes. This association aligns with previous research that demonstrated that threatening language is frequently employed alongside expressions of animosity (Roxell, 2011), as indicated by one of the strongest coefficients in the analysis. In

---

<sup>50</sup> The language used in these tweets included terms such as "sharing," "together," "love," and "refugeewelcome," which were prevalent throughout.



contexts where trigger events occur, this finding is concerning because it suggests that many posts containing anti-refugee sentiments may also contain threatening language that could potentially escalate to offline tensions and possible hate incidents. Examination of tweets with high scores for both anti-refugee attitudes and threat provides evidence to support this claim, with some advocating for direct violence to refugees or the barracks<sup>51</sup>. Results indicate that while a very small percentage of the tweet corpus contains direct threats or incitements to violence, those posts that do contain such content receive notably high scores. Additionally, a concerning trend was identified whereby many posts insinuated that the majority of threats and coordinated violent actions are carried out on closed network (Lup et al., 2015) social media platforms such as Facebook, which offer features such as closed groups:

*“Neil Hamilton it's 'appalling' there were calls in far right FB group for Asylum seekers at Penally to be shot, gassed & run over. It's 'appalling' that you recently made video with Swansea based far right activist who manages that FB group. PS. 'essential travel only' in Wales!  
<https://t.co/i2BOe5bau4--> **@NoToHateInWales***

While open network platforms like Twitter are often used to propagate general intolerant views, more extreme and coordinated actions, which have the potential to lead to coordinated offline approaches, can frequently occur in closed network spaces. This trend has been observed in other social contexts, such as the Arab Spring and London riots (Howard et al., 2011; Fuchs, 2012). A significant concern arising from this trend is that a hidden underbelly may be operating out of plain sight in the wake of trigger events, allowing radical views to be amplified in echo chambers of like-minded individuals in the network society who are connected not by spatial or temporal dimensions (Castells, 2009), but by shared anti-refugee perceptions. Related literature has suggested that closed networks like Facebook groups may amplify polarisation, potentially resulting in real-world coordinated incidents (Harel et al., 2020).

---

<sup>51</sup> Due to the Social Data Science Lab's ethical guidelines, tweets authored by individual users with high threat levels could not be published without gaining consent.

While the prevailing scholarly view asserts that such phenomena are more frequently observed in closed network structures, the findings indicate that open-network platforms like Twitter may still provide fertile ground for developing divisive echo chambers. H3 was formulated to investigate discussion networks' role in producing and propagating anti-refugee attitudes. The study employed two methods to measure discussion networks, namely, (i) examining the mention network visualisation nodes and edges and (ii) interpreting the influence of mentions in the anti-refugee attitudes regression model. The mention network visualisation showed that the online activity regarding the Penally asylum accommodation followed a core-periphery structure, which is typical of closed network or issue-specific platforms like Facebook or Reddit when a particular topic is discussed by an active community of social actors (Yang et al., 2018).

Filtering the anti-refugee posts revealed that the most central nodes in the discussion networks were the propagators of anti-refugee attitudes, while the remaining digital public dominated the less connected areas. Although visual, these findings are significant as they demonstrate the existence of extreme cases of core-periphery among the propagators of anti-refugee attitudes, indicating that discussion networks manifesting in a close-knit "echo-chamber" are a crucial component in the development and sustenance of anti-refugee sentiments, even on open network platforms such as Twitter (supporting Oxalp et al., 2020). The regression model further supported the significance of discussion networks, as mentions were predictive of anti-refugee attitudes. Additionally, both anti-refugee activists and far-right political agents had a higher usage of mentions than all remaining agent types<sup>52</sup>, which indicates that mentions are instrumental for anti-refugee content producing agents. However, as noted by Ozalp et al. (2020), the development of close-knit echo-chambers may radicalise a select few within them but in turn will receive limited public endorsement as they become out of touch with public perceptions. Public endorsement of such agents will be further discussed later in this section.

---

<sup>52</sup> All pairwise comparisons with remaining agents were statistically significant ( $<.01$ ) in the post hoc tests.

The relationship between hashtags and anti-refugee attitudes was explored in the first subset of social factors, which examined pro and anti-refugee hashtags. In relation to the reference category of 'no-hashtag', both anti-refugee hashtags and pro-refugee hashtags were more likely to involve some form of identity attack. Surprisingly, tweets with pro-refugee hashtags were more associated with anti-refugee attitudes than those with no hashtags. Although some tweets containing pro-refugee hashtags used negative language against anti-refugee advocates, these cases were not frequent enough to justify the coefficient return in the regression model. A deeper inspection of these cases revealed that there were two greater factors at play for the seemingly high level returned in tweets containing pro-refugee hashtags. First, many tweets using pro-refugee hashtags quoted tropes and arguments made online that they visibly disagreed with. Second, there were also issues of double negatives which were not detected by the classifier. For example, the prevalent pro-refugee hashtag "#nohateinwales" received a higher identity attack score due to the use of the word "hate".

This finding is consistent with previous critiques of the Perspective API classifier. The classifier is vulnerable to false alarms and detects words more generally associated with attacks on identity, rather than specifically identifying offensive terms targeted towards specific minority groups (Hosseini et al., 2017). As a result, the Perspective API classifier may be less reliable than hand-coded approaches (Kumar et al., 2021). Furthermore, this finding indicates that although Perspective API was mostly accurate when classifying anti-refugee content, it still has major limitations such as not detecting double negatives that must be explored before making empirical-driven claims. Nonetheless, tweets containing anti-refugee hashtags were extremely predictive of anti-refugee attitudes. Although the most prevalent hashtag in the dataset was pro-refugee "#LovePenallyHateRacism", the next four were all anti-refugee: "#Illegallmmigrants", "#RefugeesNotWelcome", "#MigrantInvasion", and "#Invasion". This, coupled with both the general hashtag and anti-hashtag regression output, indicates that hashtags were an important tool used to reinforce and spread anti-refugee attitudes, making it more discoverable and allowing for a contagion effect (Sunstein, 2017) of divisive sentiments in a more centralised and easy-to-access way (Yang and Counts 2010).

#### *5.4.1.2. Agent comparisons of anti and pro refugee speech*

Prior to exploring different types of social actor, general themes and motivation factors for high anti-refugee content are considered. High scoring anti-refugee tweets were examined. Many of these were either authored by individuals in the 'digital public' reference category, or by individuals in the anti-refugee activists. As such, exact extracts cannot be included, because of Hate Lab ethical guidelines. Stephan et al (2008) states that tangible or cultural factors can cause perceived ingroup threat towards the outgroup. In this context the ingroup relates to local citizens in the Penally/ Tenby area, with the outgroup referring to the camp refugees. Tangible threats can be physical harms or economic threats such as the use of public resources, while social factors can be seen as a reduction of cultural values by outgroups. Physical threats refer to the perceived risks of physical harm to the ingroup due to the presence of the outgroup. In the examined tweets, there were various instances where this threat is alluded to. For example, one tweet suggests that Penally houses "rapists,", with others referring to them as "" terrorists" or "criminals" implying perceived a danger to the local population. Another tweet questions why there was an ambulance at the camp, suggesting that there may be some sort of violence occurring at the camps. These tweets illustrate how the perception of physical harm can contribute to feelings of threat and ultimately lead to increased negative attitudes towards the outgroup.

Conversely, economic threats relate to the perceptions of the outgroup straining public resources, jobs, or other economic opportunities valued by the ingroup. In the tweets, numerous references exist to the strain placed on public resources, such as the police force and the NHS. One tweet suggests that the people of Penally cannot access the police because the resources are being used to support the refugees. Another tweet implies that the refugees are receiving preferential treatment over homeless veterans who are also in need of support. These tweets illustrate how the perception of economic strain can lead to negative attitudes towards the outgroup, as the ingroup may feel that their own economic opportunities access to resources are being threatened. Finally, cultural threats refer to the perception that the outgroup is diminishing or eroding the cultural values of the ingroup, one tweet suggests that the presence of the refugees in the community is an attack on Welsh

culture. These tweets illustrate how the perception of cultural threat can contribute to feelings of threat, leading to negative attitudes towards the outgroup. These different threats and strains can be identified as causal factors for prejudice justification, according to the JSM proposed by Crandal and Eshleman (2003). JSM posits that perpetrators of bigoted attitudes or activities do so as a result of an ongoing psycho-social conflict between "a desire to express an emotion and, at the same time, maintain values and self-concepts that conflict with prejudice" (Crandal and Eshleman, 2003: 414). In the case of the Penally asylum accommodation, the feelings of perceived threat and strain attributed to the outgroup serve as initial justifications for ingroup members to exhibit negative behaviours. In conclusion, the data collected from the tweets illustrates how different threats and strains, such as physical, economic, and cultural, can contribute to increased negative attitudes and behaviours towards outgroups.

In the second dimension of social factors, (anti-refugee content) regression model included six agent types<sup>53</sup>. Prior research has indicated that extremist political actors have played significant roles in promoting tensions in both online and offline contexts following triggering events (Legewie, 2013; Williams & Burnap, 2016; Ozalp et al., 2020). However, these studies merge multiple extremist political actors into one, often grouping official political institutions with general activists and far-right organisations. While these findings provided important insights into the primary drivers of tensions following trigger events, they failed to offer this key separation. This study builds on this by distinguishing between two subcategories: far-right political actors and anti-refugee activists. The findings provide new evidence that both subgroups are more likely to promote anti-refugee attitudes compared to the broader "digital public," but to different degrees. Notably, the results demonstrate that anti-refugee activists (+0.11) are significantly more associated with anti-refugee attitudes than far-right political actors (+0.03). This can be further explored by

---

<sup>53</sup> The coefficients derived from this model indicate the probability of anti-refugee attitudes compared to the general digital population. A positive coefficient implies a higher level of anti-refugee attitudes, whereas a negative coefficient suggests a lower level.

considering aggregate qualitative output of both groups (Figures 25 and 26), showing clear thematic differences between the language used.

Cohen's (2002) notion of "symbols of trouble" posits that certain social demographics can become predictive symbols of violence due to their demographic characteristics such as race or religion. Cohen (2002) suggests that such groups can be conflated with these symbols by the media or other institutions. Williams and Burnap (2016) tested this theory in a trigger event context following the Woolwich terrorist attacks. They found that far-right political actors were key contributors to the further conflation of religious and racial minorities with terrorist attacks and extremist behaviours. While this finding was perhaps expected given the trigger event related to a terrorist attack, it is still a concerning demonstration of predictive symbols in action. As previously noted, the empirical data reveals that far-right political actors are more prone to producing anti-refugee tweets than the general digital public.

"**sos** In light of Europe's recent record of terrorist attacks, the UK Government's decision to house 250 illegal migrants in #Penally is baffling writes @NeilUKIP <https://t.co/WcTjji93hu>": **UKIP**



Figure 23: Word cloud for anti-refugee attitudes (0.7+) tweets published by far-right political agents (top 40 words, with numbers, special characters and stop words removed)

Figure 23 reveals the conflation of refugees with terrorism (i.e. 'extremist', 'terrorism', and 'terrorist') is still prevalent, a finding that is particularly troubling given that the trigger event did not relate to a specific terrorist attack. In addition, #ISIS was a reoccurring hashtag. These results suggest that minority groups continue to be portrayed as predictive symbols (Cohen, 2002) of extremism, despite no directly related terror incident. Williams and Burnap (2016) concluded that these sentiments can be galvanised and spread when propagated online amongst many other groups, giving way to contagion effect (Sunstein, 2017). The extent to which such sentiments and claims were endorsed or condemned is discussed later in this chapter. As previously stated, this study provides novel insights of the involvement of political actors, including extremist organisations and activists, not affiliated with established political parties, in promoting hatred subsequent to triggering events. The aggregated word clouds depicted in Figure 24 visually represent the prevalent themes underlying the dissemination of anti-refugee sentiments by anti-refugee activists who are not affiliated to political parties.



Figure 24. Word cloud for anti-refugee attitudes (0.7+) tweets published by anti-refugee activists (top 40 words, with numbers, special characters and stop words removed)

This further corroborates Cohen's (2002) theory of predictive symbols, which suggests that certain terms or images can predictably evoke strong emotional responses among specific social groups. Again, the conflation of refugees with terrorism is a prevalent theme, with the word "terrorist" featuring prominently.

However, the incidence of terrorist-related language is overshadowed by other negative symbolic themes that were absent in tweets authored by far-right political agents. Anti-refugee activists commonly conflate refugees with paedophilia, utilising terms such as "children," "groom," "rape," and "rapist" with high-frequency. In most cases, these terms are used to suggest that refugees groom and sexually abuse children. An example of these claims made in context by an anti-refugee activist is given below:

*"In regards to #penally #tenby. Its done now. The army barracks will be used as an aslyum processing centre. IF, in time there is a violent rape, sexual assault, assault, robberies, murder, or god forbid a terrorist attack. @Simonhartmp and @pritipatel, it will lie at your feet!"*  
**@ChangePolitics5**

These more extreme cases of anti-refugee content can be explained by Festinger's (1952) theory of deindividuation, where anonymity reduces moral responsibility and leads to disinhibition (Joinson, 1998). This empowers offenders like anti-refugee activists, who predominately use aliases, to engage in negative behaviour and severe offenses. They target refugees as an "outgroup" (Gerstenfeld, 2017), vilifying them with unsubstantiated claims like labelling them "terrorists" or "rapists."

#### *5.4.1.3. Pro Refugee Actors – External Influencers of Collective Efficacy*

None of the remaining agent types were discovered to possess positive associations with anti-refugee attitudes compared to the digital public (reference category). As expected, few pro-refugee activist tweets exhibited even moderate levels of anti-refugee sentiment<sup>54</sup>.

---

<sup>54</sup> Only one tweet published by pro-refugee activists constituted as moderate levels (<0.6 - n=1), and in this case the tweet was quoting another user. No pro-refugee activists' tweets constituted as being severe (<0.7 - n= 0) hate.





willingness among capable and reliable actors motivated to intervene for the common good. Such interventions involve direct actions social actors take to challenge or counteract damaging behaviours in social settings. An examination of pro-refugee activist tweets revealed numerous instances of direct challenges to anti-refugee attitudes.

As suggested earlier, pro-refugee activists perhaps play a role in promoting collective efficacy by challenging and discrediting statements founded on false premises and discriminatory generalisations, which were not pertinent to the actual camps. By doing so, seek to debunk labels such as "illegal," "terrorists," or "rapists" that were often unfairly attributed to refugees, signalling such tropes as unacceptable and incorrect to wider digital audiences (Groff, 2015). This perhaps serves as a suppressive force among present and prospective offenders (Crandal and Eshleman, 2005: 245). Furthermore, they also leveraged external controls (Sampson et al., 1997), such as on-platform report functions, to mobilise users in regulating or removing far-right activists from digital spaces:

*"#penally please report far right activists @VoiceOfWales  
<https://t.co/Tzen72ByDK>": @patanna*

In addition to their condemnation of online anti-refugee attitudes, pro-refugee activists also confronted incidents that occurred offline. This suggests that interventions to maintain collective efficacy can involve actions taken in both online and offline contexts, which are not mutually exclusive. This finding reinforces the evidence related to H1, which suggests that trigger events impact both the online and offline spheres, and that positive counter-speech can likewise be observed in tandem in both contexts. This is evidenced in two ways. First is the condemnation of offline anti-refugee incidents.

*"Alarming reports tonight that far right activists planning repeat tomorrow night of shocking lynch mob video they live streamed from Penally on Monday & that leader of neo fascist For Britain party, Anne Marie Waters, will be present. See our FB post for more  
<https://t.co/5edCPDDeuY> <https://t.co/d8uCtVEk0Z>":  
@NoToHateInWales*

Notably, a considerable number of condemnations of offline incidents were redirected to other platforms, such as Facebook. One possible explanation is Facebook does not impose any character limits on text, unlike Twitter (Han et al., 2019), which allows for more detailed and nuanced discussions and critiques of anti-refugee incidents. Figure 26 provides visual evidence of this depicting a video recording of a particular incident accompanied by a lengthy discussion featuring crucial details.

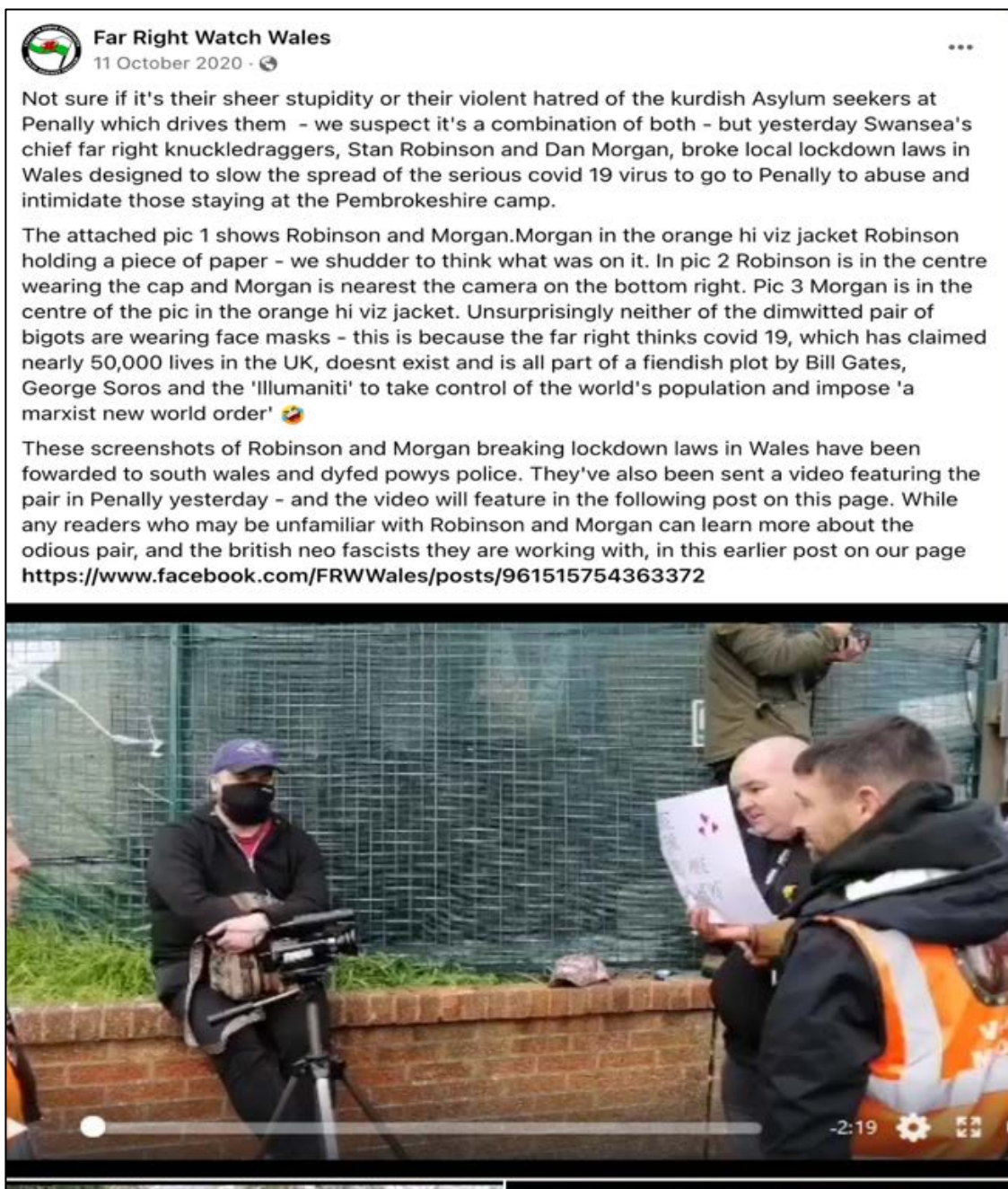


Figure 26. Pro-refugee activist Facebook post

Another approach employed by pro-refugee activists to advance collective efficacy in both online and offline spheres is the mobilisation of offline actions, including but not limited to protests and welcoming activities. These endeavours are not solely directed towards addressing specific incidents but are intended to engender support for refugees in a more comprehensive manner. Excerpts provided below illustrate some of these activities.

*“This morning in Penally crowds of anti-racist protesters have gathered again at the gates of an army base housing asylum seekers. They are trying to stop far right and fascist groups attacking the base. Read our full report here: <https://t.co/gX59aGoCCS> [https://t.co/JO7ySHe7cU\\_](https://t.co/JO7ySHe7cU_): @voice\_wales*

*“PENALLY WELCOME PACKS. Please donate! [https://t.co/yV8jqhpb9D\\_](https://t.co/yV8jqhpb9D_): @sutrvale*

Police, political and news agents were all less associated with anti-refugee speech than the digital public. Extracts reflecting the tweets produced for each agent type are given below<sup>55</sup>:

*“Superintendent Anthony Evans, Divisional Commander for Pembrokeshire, said: “A peaceful protest was held last night with around 20 people in attendance and minimal disruption. We will maintain a presence in Penally to facilitate peaceful protests and continue to engage with the public”: **Heddlu Dyfed-Powys Police (ArAS: 0.03)***

*“Police were called to the Penally Asylum Accommodation centre on Tuesday afternoon: <https://t.co/6GAGCbzAUf>”: **BBC News (ArAS: 0.07)***

*“We welcomed today’s four nations meeting to discuss how to help asylum seekers and wider communities to integrate more effectively in Wales and the UK” **Welsh Government (ArAS: 0.11)***

---

<sup>55</sup> ArAS= Anti-refugee attitude scale score from perspective API classifier

This underscores the expectation that language used by police, news outlets, and political figures is characterised by its brevity, precision, impartiality, and, above all, the absence of anti-refugee language. Using objective and unbiased language creates an impression of reliability, a crucial prerequisite for fostering collective efficacy while mitigating tensions (Sampson et al., 1997). Although various entities may use social media to counteract negative attitudes directed at refugees, Sampson (1997, p. 918) posits that the collective efficacy effect can only be achieved if the public harbours a degree of tolerance and trust. This trust is often built upon by established and trusted institutions adopting impartial and objective stances following such incidents so to not fan any flames (Ozalp et al., 2020). These findings, as anticipated, suggest that influential and trustworthy social actors were more inclined to discuss the asylum accommodation incident using neutral and factual language, thus avoiding adding to the tensions observed in the general public discourse. Consequently, these institutions did not contribute to the proliferation of anti-refugee sentiments, a contrast to instances where certain institutions, like the media, have at times exacerbated negative perceptions (Williams and Burnap, 2016). However, these findings are limited in their ability to determine the effect such sentiments may have had on the general public. This aspect will be further discussed in the subsequent sections.

In addition to news agents not propagating anti-refugee sentiments, neither of the two external factor predictors, google searchers and press headlines, returned any significant associations. This is particularly surprising, given previous evidence supporting Cohen's (2002) proposition that traditional news sources are instrumental in agenda-setting and constructing narratives that perpetuate tensions (Williams and Burnap, 2016). Although Google searchers and media outlets may play a pivotal role in the construction of divisive narratives in large-scale violence-oriented trigger events, evidence in this study suggests a diminished role in localised and non-violent policy-driven events, such as the establishment of the Penally asylum accommodation. The trigger event in question, not being linked to acts of violence such as terrorist attacks, may have prevented the media (headlines) from labelling refugees as "predictive symbols of violence" (Cohen, 2002), as has been observed in prior studies (Legewie 2013; Hanes and Machin 2014; King and Sutton 2014; Williams and Burnap, 2016) to instigate anti-refugee attitudes. In response, far-right

political agents and anti-refugee activists adopted a more direct approach (on Twitter), labelling refugees as symbols of violence, despite no recorded incidence of violence by them. Previous research has shown that the media can "fan the flames" of tensions after extremist events (Williams and Burnap, 2016). However, in the absence of established institutions playing this role, far-right agents seem to have assumed the entire responsibility, propagating such sentiments through more direct channels such as social media.

#### ***5.4.2. Public Response to Online Discourse: Endorsements and Counters***

In H4, the extent to which online discourse relating to Penally is endorsed, propagated, or challenged was examined. This was measured by exploring how and the extent to which tweets are interacted with online. In line with related work, anti-refugee attitudes was not propagated by retweets (Williams and Burnap, 2016; Ozalp et al., 2020). This suggests limited endorsement among the general public and serves as an important initial indicator of the existence of a digital collective efficacy (Sampson et al., 1997). Far-right political agents were negatively associated with all three-engagement metrics, indicating that they receive extremely limited public endorsement across the board. However, a novel finding made possible by including additional engagement metrics was that anti-refugee content was positively associated with likes and replies. This was likewise the case with anti-refugee activists, who were negatively associated with retweets, indicating that the digital public were unlikely to propagate the content, but conversely may still like or reply to it. The positive association between anti-refugee content (and anti-refugee activists) with likes presents some concern. Perhaps many may endorse some of the anti-refugee sentiments but do not wish to appear to be actively supporting them by retweeting.

As mentioned earlier, a close-knit echo chamber was discovered of a small number of social actors being responsible for the majority of anti-refugee content. The positive association with likes suggests that a silent group of supporters may exist who share the same sentiments towards refugees as the dense core of anti-refugee

content but are less vocal in making or supporting them. These findings were undetected by previous studies, such as Williams and Burnap (2016), which only measured endorsement by more active propagation measures such as retweeting, indicating that many views may still be supported but in more passive ways such as likes. In this particular context, the prevalence of collective efficacy is assumed to be higher when evaluating what Warner (2014) classified as active forms. This involves the public's abstention from actively endorsing anti-refugee content by disseminating it through retweets. However, passive forms of endorsement, such as likes may be more substantial than previously anticipated. A positive finding was that threatening content was negatively associated with all three-engagement metrics, indicating extremely low levels of endorsement for any content that used threatening language.

Initially, the associations between anti-refugee attitudes and anti-refugee activists with replies are of concern, as they were strongest associations with replies in the model. However, a closer examination shows replies to anti-refugee tweets, shows that these correlations may not necessarily indicate public endorsement, but instead, pro-refugee counter-speech. This underscores the pivotal role of community members themselves in utilising counterspeech as an effective means of "combating hate" and "fostering peaceful, non-polarised discussions," (as indicated by Garland et al., 2022:1).

*"They are migrants, please reveal your source that they are here 'illegally'": **Tom***

*"We don't need UKIP in Pembrokeshire. We don't want UKIP in Wales. In fact, we don't want UKIP at all @YesCymru @Plaid\_Cymru": **Susan***

*"If only you could turn your attention to something vastly more important than stirring hatred against your fellow human beings": **Jack***

The extracts<sup>56</sup> demonstrate that most replies to anti-refugee content came from members of the digital public who actively challenged the views expressed. For

---

<sup>56</sup> Permission was obtained for all three extracts to be included, albeit aliases were employed to maintain anonymity.

example, one user challenged the notion that refugees are "illegal." These findings support the idea that collective efficacy is more effective and common in active forms of engagement, such as challenging or intervening with antagonistic speech and agents (Simons et al., 2005). This is perhaps not surprising, as Sampson (2001) suggests that collective efficacy is primarily driven by capable and willing actors who actively intervene in negative social behaviours. In the JSM Crandal and Eshleman (2003) assert that interventions aimed at curbing antagonistic speech serve as potent suppression forces by publicly denouncing these views and reinforcing social norms that prohibit it. These interventions are significant as they effectively counterbalance the justifications previously outlined. While responding to this content is vital, such efforts are often situational. It is thus imperative to consider the endorsement of such messages by proactive social actors such as pro-refugee activists, who intervene in a more sustained and pre-emptive manner. Consistent with prior research, this study found a positive association between pro-refugee activists and information propagation in the retweets model (Ozalp et al., 2020).

Additionally, pro-refugee activists were positively associated with likes, (a stronger association compared to anti-refugee activists). These results suggest that while passive support for anti-refugee sentiment may exist, it remains lower than the passive support of counter-movements. In analysing the positive association between pro refugee activists and replies, caution was exercised in attributing the positive association to general endorsement, as it may reflect counter-speech, as seen with anti-refugee content. However, after examining numerous replies, no counter-speech was detected, and almost all comments expressed support and solidarity. These findings underscore the importance of organisations dedicated to safeguarding communities and enhancing collective efficacy on social media (Ozalp et al., 2020). Furthermore, tweets that employed positive sentiment language supportive of refugees were positively associated with likes and retweets, indicating widespread endorsement of the general notion of supporting refugees. This further supports the notion that positive counter-movements are more widely endorsed by the digital public.

Other agent types exhibited negative relationships with anti-refugee content, including media, news, and political agents. These established institutions reflect



what Sampson (2001) described as external forces to collective efficacy. While he acknowledged that informal controls are most salient, he maintained that external forces still play a significant role in upholding collective efficacy, particularly in times of great strain, such as trigger events. Their role can be better understood by considering validation theory (Edwards and Rushin, 2018), that posits that established institutions that are perceived as trustworthy and legitimate can play a significant role in influencing information flows (Sampson, 2001). The extent to which they are perceived as legitimate and trustworthy is described as symbolic capital (Dupont, 2004) and can vary significantly. Previous work has outlined validation theory as an amplifier for dangerous rhetoric, when a perceived legitimate force, such as the media<sup>57</sup> or political agents uses their symbolic capital to disseminate negative information that can be discriminatory towards minority groups (Williams and Burnap, 2016). Despite this, results in this chapter show that none of the police, political nor media were associated with producing anti-refugee content. This presents an opportunity for these agents to use symbolic capital to contribute to collective efficacy by propagating unbiased content.

The study found a positive association between political agents and all three-engagement metrics, which aligns with previous research that has highlighted the role of political legitimisation in driving support for policies or issues (Edwards and Rushin, 2018). This indicates that political agents' symbolic capital and authority make them effective communicators and influencers in shaping public opinion. A notable finding in this study is that political agents were negatively associated with antagonistic content, implying that they did not use their symbolic capital to legitimise or validate these views. This contrasts with other cases in which political agents with the legitimacy to propagate radical views have legitimised dangerous rhetoric (Perry, 2001). The neutral language used by political agents in this study indicates that they were not using their platform to promote negative attitudes towards refugees. The digital public's widespread endorsement of predominantly positive content propagated by political agents provides additional evidence of the importance of community protection agencies in curtailing anti-refugee attitudes (Ozalp, 2020). The

---

<sup>57</sup> The media was predominately measured by exploring written news headlines and readers should note other media forms could exhibit notable differences.

study's findings highlight the crucial role that political agents play as key players in shaping public opinion and attitudes towards refugees. Their ability to generate engagement through positive messaging underscores their potential impact in promoting inclusive and welcoming attitudes towards refugees. These findings contrast sharply with the results for far-right political agents, who were positively associated with anti-refugee content but negatively associated with all three-engagement metrics. This highlights that although both political and far-right political agents possess symbolic capital derived from political legitimisation, anti-refugee sentiments are not endorsed, lending additional support to the presence collective efficacy (Sampson et al., 1997).

Based on the analysis of police Twitter accounts, there is a high association between the police and the number of replies to their tweets, but no significant association with the number of likes or retweets received. This suggests that the police are actively engaging in discussions with their Twitter followers, but they are not garnering as much attention in terms of likes or retweets. Upon closer inspection of the replies to police tweets, it became evident that social media is used as a platform to maintain open discussion networks between the police and the public. This enables the public to receive transparent updates on the asylum accommodation while also allowing for information collection from the public. Such open communication networks between the police and the public on social media can facilitate information-sharing, leading to improved outcomes for all involved (Mawby, 2002). As such, social media platforms like Twitter can play a valuable role in developing and maintaining transparent communication networks between the police and the public.

Press headlines were positively associated with all three metrics. This finding supports the classic criminological notion that the media continues to have a significant role in shaping public discourse following significant events by "setting the agenda" and "transmitting the images" (Cohen, 2002), even in a more localised case, through traditional approaches such as publishing stories in papers. Media outlets were not associated with anti-refugee attitudes, suggesting that they were not fuelling such sentiments as seen in other wider-scale events related to extremist events or political votes (Hanes and Machin, 2014; King and Sutton, 2014; Moore

and Ramsey, 2017). These results are encouraging as they demonstrate that the media could play a more neutral role in local policy events while still retaining influence over information flows. Therefore, when reporting accurately and unbiasedly, they can serve to debunk misinformation. They may contribute to collective efficacy by using their symbolic capital to challenge negative attitudes, reinforce intergroup trust, promote mutual understanding, and denormalise radical opinions that drive community tensions (Sampson et al., 1997).

## 5.5. Summary

This chapter presented further evidence supporting computational criminology methods such as data aggregation, machine classifiers and statistical modelling for detecting, measuring, and understanding information flows and the production of anti-refugee content following a trigger event. Additionally, this chapter makes a methodological contribution by demonstrating that the Google Perspective API, which detects identity attacks, can serve as a viable alternative to in-house machine classifiers employed by public bodies and academic institutions like Burnap et al. (2015) and Ozalp et al. (2020) for analysing trigger events. This approach offers advantages such as greater accessibility to researchers with limited access to computer scientists, faster and more cost-effective analysis, and the potential for a standardised methodology that can be replicated and compared across various trigger event case studies. While the methodology reports good accuracy of the Perspective API, some misclassifications were also observed, indicating that the Perspective API is not entirely reliable. This issue will be further explored in subsequent sections.

From a knowledge perspective, this chapter analyses the factors influencing public discourse and attitudes towards refugees in response to a localised trigger event in Wales. The case study examines the Penally asylum accommodation, where an influx of refugees disrupted a previously culturally homogenous area, primarily inhabited by a white population. This investigation deepens our understanding of trigger events, particularly within these types of communities, and provides valuable insights. This observation aligns with Garland & Chakraborti's (2006) research,

which suggests that within close-knit communities characterised by a strong collective identity, certain members may contribute to rejecting minority groups based on racial distinctions, leading to heightened tensions. It underscores that hidden biases can exist even in seemingly harmonious communities, emphasising the need for a nuanced discussion on these dynamics. The study examines the announcement of the asylum accommodation as the trigger event and their duration and eventual closure to shed light on the factors that shaped public perceptions of the refugee population in this specific context. Notably, this investigation leverages emerging digital techniques that offer fine-grained temporal analysis of locomotive social media data, supported by computational criminological methods. These methods build upon similar studies that have explored the relationship between antecedent trigger events and observable public social media reactions (Williams and Burnap, 2016; Ozalp et al., 2020).

The findings highlight the crucial role of collective efficacy in countering antagonistic attitudes and promoting inclusive attitudes towards refugees (Sampson et al., 1997). Political and community protection agencies played a crucial role in shaping public opinion and promoting welcoming attitudes towards refugees, receiving high levels of endorsement from the public as measured by engagement metrics (likes, retweets, and replies), underscoring their potential impact in promoting inclusive attitudes. Conversely, the study found anti-refugee agents, such as far-right political and anti-refugee activists, were significant contributors to anti-refugee attitudes. Anti-refugee activists were significantly more likely to author antagonistic tweets, which could be explained by their feeling of anonymity when using aliases, linking to Festinger's (1952) theory of deindividuation. These anti-refugee tweets were more severe and detached from reality, such as calling refugees rapists and terrorists. Fortunately, anti-refugee tweets and associated authors received limited levels of public endorsement across most engagement models. Some anti-refugee attitudes was positively associated with replies, but qualitative extracts showed that these replies were predominantly counter-speech challenging the anti-refugee content published by pro-refugee actors. Furthermore, the study found a core-periphery structure (Borgatti and Everett, 2000), whereby a small minority of users contributed the most anti-refugee content. This means that although at first glance a significant amount of anti-refugee content existed, it was predominantly contained to a small echo

chamber (Ozalp et al., 2020), indicating that a targeted approach could be effective in reducing this sort of content in similar events. The most severe cases of anti-refugee attitudes were observed in the initial stages of the policy change trigger event, supporting the idea that racial tensions are most extreme directly after sudden shifts in heterogeneity (Putnam, 2007; Twigg et al., 2010).

Furthermore, this chapter highlights that digital platforms can be valuable tools for maintaining transparent communication networks between the police and the public, ultimately leading to improved outcomes for all involved with greater clarity and information provided. The study found that the media<sup>58</sup> still plays a significant role in agenda-setting and transmitting information (Cohen, 2002). However, it is important to note that in this localised event in Wales, the media did not contribute to anti-refugee sentiment, meaning they did not fan the flames but were still influential in information flows (Williams and Burnap, 2016). As such, partners should work with the media to ensure accurate information that debunks false claims propagated by other agents. Overall, this study highlights the crucial role of collective efficacy in countering antagonistic speech and promoting inclusive attitudes towards refugees (Sampson et al., 1997). The findings also emphasise the importance of prompt action and effective communication strategies to counter antagonistic speech and promote collective efficacy in such situations.

---

<sup>58</sup> The media was measured by exploring written news headlines and readers should note other media forms could exhibit notable differences.

## Chapter 6

### Promoting Cohesion Through Digital Spaces: Strategies and Approaches of the Third sector in Mitigating Tensions

---

#### 6.1. Introduction

Scholars have highlighted the unique position of the third sector in fostering collective efficacy by bridging external institutions with communities (Kelly, 2007). As hybrids of public and community institutions, NGOs can be critical in networking with and between communities (Blake et al., 2008). NGOs not only act as a pressure point for communities to engage with formal actors but can also directly contribute to collective efficacy using informal mechanisms themselves. This is generally divided into two approaches: "cohesion and trust" and "informal social control" (Sampson et al., 1997). Scholars have argued that cohesion extends to various domains, including "common values and a civic culture" (Forrest and Kearns, 2001). Social controls are based on rewards and penalties, which derive from either adhering or deviating from societal norms (Groff, 2015).

Despite the critical role of NGOs in promoting collective efficacy, little academic research has been conducted to examine the specific strategies and tactics used in these campaigns. This chapter bridges this gap by investigating how NGOs promote positive and inclusive content, particularly within digital spaces, and it seeks to measure the effectiveness of various strategies employed. Building on existing literature, which highlights social media and digital content as potent tools for fostering mutual understanding and addressing issues related to community tensions (Chakrabarti, 2018:400), it specifically identifies and explores content strategies used in this context. To accomplish this, this study utilises Mergel's (2013) framework, primarily used in the context of public sector organisations. By applying this framework to civil society this study contributes to the existing literature on the role of the third sector in promoting collective efficacy. This study also employs regression

models to assess which approaches are most effective in achieving collective efficacy through social media, providing evidence-based recommendations to NGOs.

The chapter analyses the communication strategies and sub-strategies employed by minority rights organisations through a content analysis. Four broad strategies (push, pull, networking, and transactional) are identified that reflect Mergel's (2010) communication typology, and eight sub-strategies are inductively derived within these. To do this, analysis considered content by ten prominent non-profit organisations in Wales, by analysing a sample of 5,000 tweets collected over 27 months, specifically from November 10, 2018, to February 10, 2021. Given the relatively recent adoption of social media platforms by NGOs in Wales, the study was constrained by the limited availability of a larger dataset. A concurrent mixed methods approach was employed to analyse the data, incorporating quantitative methods to identify prevalent strategies and breakdowns of attributes, alongside qualitative methods that provided contextualised examples of usage, thus enhancing the context and our understanding of the findings<sup>59</sup>. The study aims to provide evidence-based recommendations to NGOs on promoting positive, inclusive, and diverse content while combating divisive and misinformed attitudes. The findings of this study contribute to a better understanding of the role of NGOs in promoting collective efficacy and provide insights into effective communication strategies.

## 6.2. Hypotheses

*H1: NGOs will use social media to contribute to collective efficacy (Sampson et al., 1997) by using informal mechanisms, mainly promoting cohesion and trust through positive campaigns that support inclusiveness, in addition they will also sometimes*

---

<sup>59</sup> It is pertinent to acknowledge that the results of this investigation exclusively relate to NGOs that are engaged in mitigating tensions and may not be generalisable to other third sector entities, or public and private institutions. Nonetheless, the findings presented in this chapter furnish a valuable resource for organisations striving to mitigate online tensions via efficacious communication strategies.

*use informal social controls, such as actively challenging divisive attitudes with public interventions to promote collective efficacy.*

Sampson et al. (1997) propose that informal social controls and promoting cohesion and trust are the two primary mechanisms for achieving collective efficacy. While existing research on digital collective efficacy has primarily focused on the former, typically in response to specific instances of community tensions (Costello et al., 2017; Ozalp et al., 2020), also known as counter-speech (Williams and Burnap, 2016), the current study places greater emphasis on the latter. Specifically, this study examines the day-to-day actions taken by NGOs to mitigate tensions and promote inclusivity through general promotion campaigns that expose audiences to positive and inclusive content (Perry et al., 2016; Chakraborti, 2018). This approach aligns with Allport's (1954) intergroup contact theory, which suggests that positive interactions between social groups, including awareness campaigns, can foster social change, termed "positive contact". Accordingly, it is hypothesised that NGOs will employ cohesion and trust strategies more frequently than interventions. However, interventions may generate greater endorsement and engagement, particularly in response to trending incidents that attract heightened public interest and engagement.

*H2: Third sector minority rights organisations will harness social media to enhance their social capital and networking with communities and external stakeholders.*

This hypothesis is founded on existing literature that suggests social media possesses the ability to facilitate social inclusion and encourage community involvement (Vance et al., 2009). Dialogical strategies, which emphasise two-way communication and collaboration, have been the most common and effective in building social capital with community members and external partners (Bortree and Seltzer, 2009; Lovejoy and Saxton, 2012). Within the dialogical paradigm, both pull and networking strategies are commonly used by organisations to engage in direct dialogue with their audiences on social media platforms (Greenberg and MacAulay, 2009). Networking strategies, which facilitate a two-way interactive framework and allow for co-design of services, are particularly effective in building issue networks and bolstering social capital with communities (Mergel, 2013). Similarly, pull



strategies that encourage public input and participation have also been shown to be effective in fostering relationships and social capital with external stakeholders and communities (Mergel, 2012). Thus, third sector minority rights organisations are expected to employ dialogical strategies, including pull and networking approaches, on social media to enhance their social capital and networking with external stakeholders and within communities.

*H3: In some cases, NGOs will use social media as an opportunity to foster a transactional environment with communities.*

This hypothesis is grounded on the assertion that NGOs commonly employ social media to solicit donations, which has become a longstanding custom within this sector (Deschamps and McNutt, 2014). This practice distinguishes them from the public sector, where transactions generally pertain to the provision of public services to communities or constituents (Mergel, 2012). The potential for NGOs to utilise social media as a platform for cultivating a transactional environment with their communities extends beyond the mere solicitation of donations. As Mergel (2010) outlined, transactional strategies involve undertaking transactions with external stakeholders or audiences relating to the provision of physical goods or services. Hence, NGOs may employ social media to offer support services or facilitate other forms of transactions that benefit their communities. It is not assumed that NGOs extensively employ social media for transactional purposes, beyond soliciting donations, given that other types of charitable organisations that are less oriented towards support services have not been previously explored (Deschamps and McNutt, 2014).

*H4: Particular strategies will be implemented to use the public as 'sensors' for information gathering for organisations.*

According to Mergel (2012), pull strategy tweets are intended to encourage followers and stakeholders to engage and participate in providing information or general input. By leveraging public attitudes towards cohesion and tensions, such approaches have been demonstrated to enhance understanding about specific issues (Meijer and Thaens, 2013). It is anticipated that minority rights NGOs organisations will

employ similar strategies aimed to engage with their communities and adopt a ground-up approach that emphasises public attitudes towards cohesion. Such organisations traditionally elicit information through appeals for input, such as surveys, attending events, or answering questions (Cullen, 2022). Thereafter it is hypothesised that third sector minority rights organisations will utilise pull strategy tweets and other comparable strategies to interact with their communities and gather information. These organisations will use a variety of methods to collect information beyond surveys, such as focus group attendance and direct outreach to community members. Through these efforts, third sector minority rights organisations may better understand the needs and perspectives of the communities they serve, and tailor their responses to address specific issues.

### 6.3. Data Variables

Table 17. Study variables (n= 5000)

Variables	Coding	Sample	
		SD	M
<b>Dependent Variables</b>			
Like Count	Range: 0-1684	29.14	0.48
Retweet Count	Range: 0-1140	22.64	0.40
Reply Count	Range: 0-26	0.75	0.27
<b>Independent Variables</b>			
<b>Content Factors</b>			
Word Count	Range: 1-56	13.38	27.18
Hashtag No.	Range: 0-11	1.36	0.74
Mention No.	Range: 0-17	1.71	0.83
Hyperlink	0 = no; 1 = yes		81
Other Focus	0 = no; 1 = yes		7%
Mention Covid-19	0 = no; 1 = yes		10%
Intervention	0 = no; 1 = yes		5%
Celebration	0 = no; 1 = yes		4%
<b>Strategy Factors</b>			
Ref: Push	0 = no; 1 = yes		57%
Pull	0 = no; 1 = yes		29%
Networking	0 = no; 1 = yes		13%
Transactional	0 = no; 1 = yes		1%
<b>Sub-Strategy Factors</b>			
Ref: Non-Organisation Information	0 = no; 1 = yes		9%
LOC (non-social media)	0 = no; 1 = yes		13%
LOC (social media)	0 = no; 1 = yes		22%
Organisation Information	0 = no; 1 = yes		14%
Interactive	0 = no; 1 = yes		16%
Work Recruitment	0 = no; 1 = yes		3%
Observational	0 = no; 1 = yes		7%
Service Use	0 = no; 1 = yes		2%
<b>Account Factors</b>			
<i>Tweet Frequency</i>			
Ref: High Tweet Frequency	0 = no; 1 = yes		40%
Low Tweet Frequency	0 = no; 1 = yes		40%
Mid Tweet Frequency	0 = no; 1 = yes		20%
<i>Account Size</i>			
Ref: X-Large	0 = no; 1 = yes		30%
Small	0 = no; 1 = yes		20%
Medium	0 = no; 1 = yes		30%
Large	0 = no; 1 = yes		20%
<b>Hashtag and Emoji Type</b>			
<i>Emoji Type</i>			
Ref: No Emoji	0 = no; 1 = yes		89%
Activity Emoji	0 = no; 1 = yes		3%
Symbol Emoji	0 = no; 1 = yes		3%
Emotive Emoji	0 = no; 1 = yes		2%
Group Specific Emoji	0 = no; 1 = yes		2%
<i>Hashtag Type</i>			
Knowledge #	0 = no; 1 = yes		14%
Event #	0 = no; 1 = yes		6%
Time/ Place #	0 = no; 1 = yes		3%
Branding #	0 = no; 1 = yes		7%
Call-to-Action #	0 = no; 1 = yes		2%
Goals #	0 = no; 1 = yes		7%
Dialogic #	0 = no; 1 = yes		5%
<b>Controls (Time/Day)</b>			
Weekend	0 = no; 1 = yes		13%
Ref: Evening	0 = no; 1 = yes		18%
Night	0 = no; 1 = yes		2%
Morning	0 = no; 1 = yes		26%
Daytime	0 = no; 1 = yes		54%

## 6.4. Results

Table 17 presents the study variables, mean values, percentage occurrences and coding of the tweets collected for analysis in this chapter ( $n = 5000$ ). The three dependent variables correspond to Twitter engagement metrics: likes, retweets, and replies. Additionally, there are five distinct variable categories or factor groups. First, content factors pertain to tweet content, encompassing elements such as hashtag use, hyperlinks, and specific mentions, including COVID-related references. Second, strategy factors are divided according to Mergel's (2012) framework and the subsequent sub-strategies identified in this study; these are further elaborated upon later in this chapter. Third, account factors relate to the organisation's social media presence, specifically the number of followers and tweet frequency. Next, hashtags and emoji types were coded into various themes, serving distinct purposes (detailed information on the methodology, coding, and categorisation of these themes is provided in the methodology section). Finally, control variables include the time of day and the day of the week.

### 6.4.1. *Outlining Strategies*

The initial coding stage involved categorising tweets into one of four strategies belonging to Mergel's (2013) social media approach typology: push, pull, networking and transactional. A breakdown of the usage of each of Mergel's (2013) strategies with further sub-strategies is presented in Figure 27. Alongside identifying thematic differences between strategies based on textual data, this section provides preliminary insights into attribute differences between each strategy. These attributes include hashtags, mentions, word count, emoji usage, and hyperlinks. To examine these differences, a Kruskal-Wallis H test<sup>60</sup>, which produces rank-based nonparametric output indicating differences in attribute prevalence among strategies, was employed. The results for all attributes indicated strong evidence for a difference ( $p < 0.01$ ) between the mean ranks of at least one group pairing. The accompanying

---

<sup>60</sup> The full Kruskal-Wallis H test results are presented in Appendix 12

mean rank statistics<sup>61</sup> serve to indicate the rank order of attribute prevalence between tweet strategies. Attribute prevalence for each strategy is presented in the form of a radar graphs that reflects attribute rank between strategies.

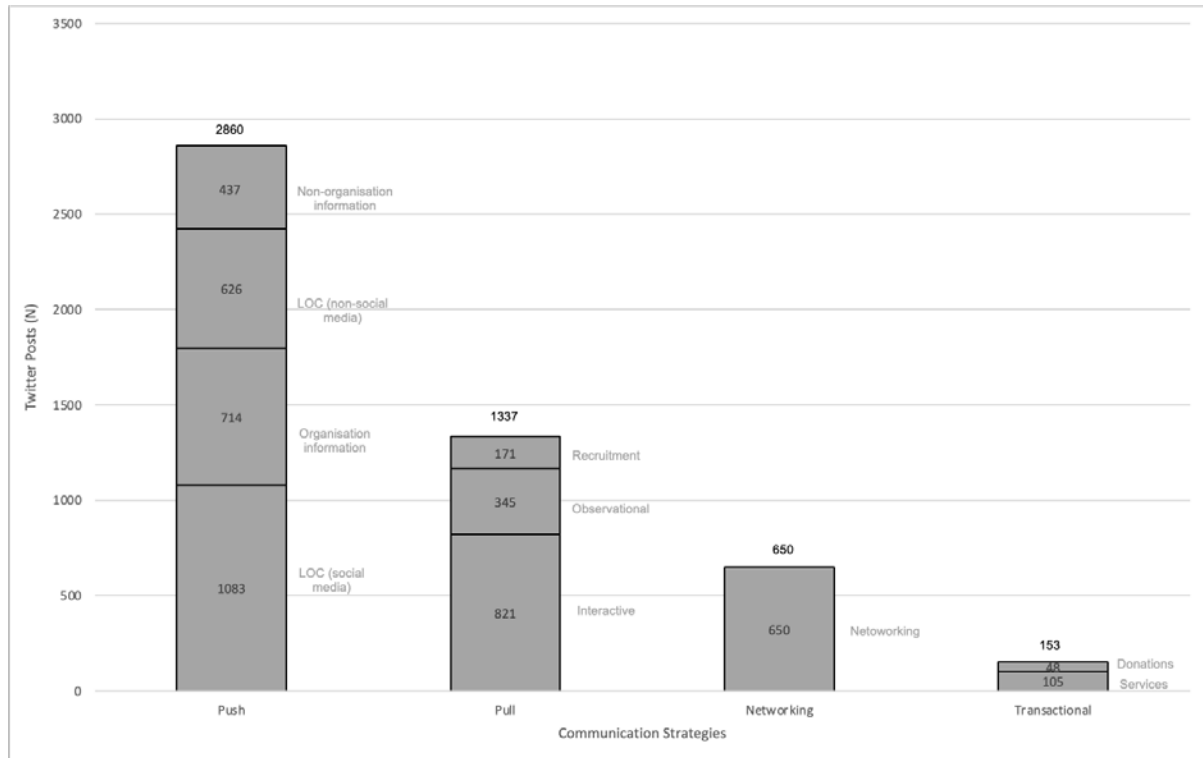


Figure 27. Tweet strategy prevalence breakdown

#### 6.4.1.1. Push Strategy

According to Mergel (2012), push strategy communications involve one-way dissemination of information by organisations, using social media as a broadcast channel to digital audiences. Such communication aims to provide content to communities and stakeholders, rather than directly engaging with them online, thus representing an informational approach within a broadcast paradigm (Greenberg and MacAulay, 2009). Social media allows organisations to broadcast information at any time, without conforming to more traditional processes i.e. through the mainstream

<sup>61</sup> Mean ranks should be interpreted cautiously because specific mean differences between strategies may not hold overall significance.

media (Lee and McGovern, 2014; Goldsmith, 2015). This is important for minority rights organisations whose content may not always be deemed newsworthy by the media. This study, which applies Mergel's (2010) framework to NGOs, reveals that informational content can cover various topics at local, regional, national, and international levels. Push strategy tweets were the most common approach and accounted for 57.2% (n=2860) of tweets, which reflects high prevalence's reported in the private sector (O'Connor, 2015). This has often been attributed to the desire to maintain control over the published content and the familiarity of broadcast approaches from earlier internet iterations, such as web 1.0 (Greenberg and MacAulay, 2009; Heverin and Zach, 2010). Despite this, a further examination of push approaches demonstrate that although they are one-way in nature, many have integrated new facets that extend beyond just broadcasting information. Such approaches are identified by considering further sub-strategies. The sub-strategies include (1) organisation information, (2) link to other content (non-social media), (3) link to other content (social media) and (4) non-organisation information:

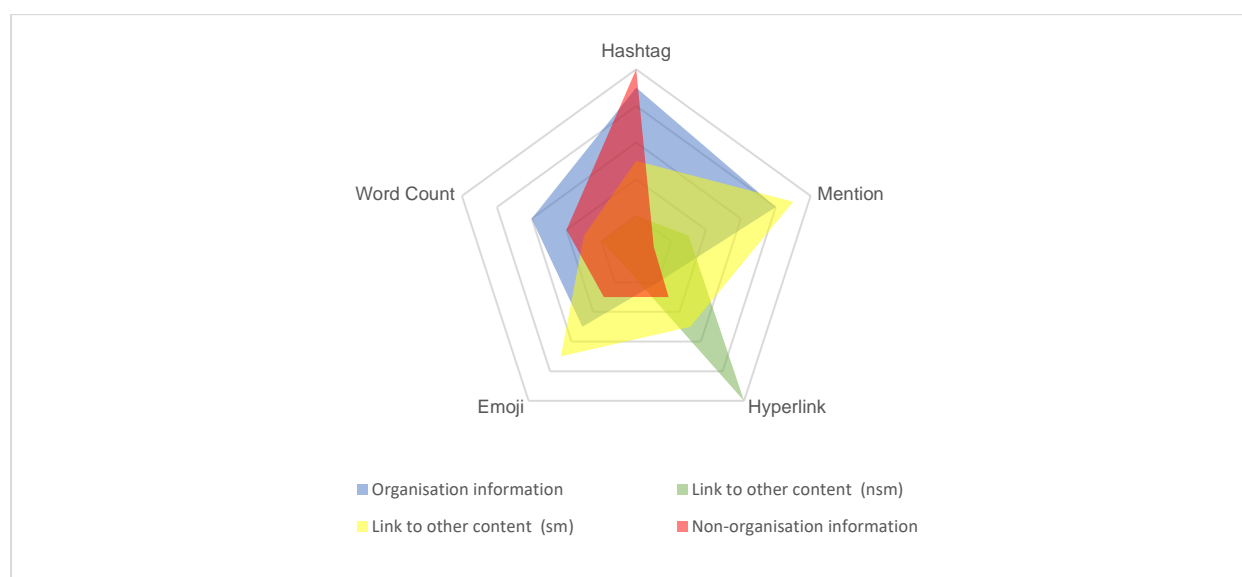


Figure 28. Push strategy attribute radar graph

#### 6.4.1.2. Organisation information

Organisation information tweets accounted for 25% (n=714) of all push strategy tweets and 14.3% of all tweets in the study. This strategy informs readers of the

activities and work carried out by the NGOs. In the public sector, organisation-specific content dominates one-way communications. It has frequently been attributed to organisations engaging in image building processes to improve public perceptions by enhancing transparency, accountability, and trust (Mawby, 2010). Image building usually involves outlining organisation successes (Kappeler and Gaines, 2020) and providing transparency about the day-to-day operations (Mawby, 2010). Findings for NGOs reflected this, with reported activities such as attendance of meetings with other stakeholders, attendance of protests, changes to the organisation itself, additional funding received or given, and summaries of work carried out by the organisation. Organisational tweets allow organisations to provide transparent accounts of their operations to online audiences, allowing them to build an online presence in real or near real time (Van De Velde et al., 2015). Examples of these are given below:

*"We've supported over 200 new refugees so far this year! Our services reducing homelessness and help refugees to build new futures in Wales. Something to be proud of! #PreventingHomelessness #ReducingDestitution #StayingSafe" --- (Welsh Refugee Council)*

*"Some AWESOME creations here made by families in Wrexham, using art packs provided by @Arts\_Connection. Distributed through our #BMECYP project with help from People for People in the region #art #puppetry #youthwork #play <https://t.co/EXXxEtChJn>" -- (EYST)*

*"We look forward to bringing our knowledge and experience towards working together with the Welsh Government and other public and voluntary agencies to ensure that the rights of our trans siblings are recognised and realised and that they are not rescinded"--- (Pride Cymru)*

*"Thank you to our Board Member @VinceDonovan from @Safety\_Sol\_ for this great training session yesterday. You showed us some amazing tips and features which we were all surprised even existed. A really engaging, informative and enjoyable session"--- (Race Equality First)*

Tweets related to organisational strategy showed high usage of hashtags (9/10) and mentions (8/10). The prevalence of hashtags was expected, as they have been

previously identified as effective tools for building public images and increasing searchability (Zappavigna, 2011). Most hashtags used were specific to the organisation, as evidenced in the second extract. In contrast, the high usage of mentions was surprising, as organisational approaches are typically viewed as one-way forms of communication. Previous studies have attributed higher usage of mentions to image building, in which day-to-day operations are discussed and efforts are made to reinforce existing ties and establish new ones, rather than simply disseminating information (O'Connor, 2015). This is also evident in the second extract, which discusses organisational activities while promoting partner organisations. Another potential explanation for the high usage of mentions is that they highlight individual members of the organisation who participate in its operations, thereby increasing transparency. Hyperlink usage was among the lowest (2/10), which is unsurprising given that organisational content typically consists of one-way broadcasts on social media platforms, making it unlikely that audiences will be directed to other websites or social media platforms (Mergel, 2012).

#### *6.4.1.3. Non-Organisation Information*

Evidence suggest that public sector organisations predominantly communicate about organisation-specific content or content that directly interacts with organisations (Meijer and Thaens, 2013). In contrast, third sector approaches show a greater willingness to broadcast information unrelated to the organisation. Non-organisation information strategy tweets accounted for 15.3% (n=437) of push strategy tweets and 8.7% of all observed tweets. These tweets were identified by considering three dimensions: (1) does not reference any work or activities carried out by the organisation (2) does not quote, mention, or reply to other accounts or tweets, and (3) does not contain any links to further organisation websites or social media pages. Non-organisation information tweets were primarily used to discuss, identify, condemn, or celebrate events that are autonomous to the work directly carried out by the account's organisation, such as religious events, news stories, key statistics or facts, and national/international days of celebration. This sub-strategy allows organisations to promote content that provides positive knowledge and awareness of minority groups, debunk misinformation and challenges bias attitudes directly,



leading to greater understanding and acceptance of minority groups (Chakraborti, 2018). This approach is closely related to early conceptualisations of positive contact theory. It posits that positive interactions with communities or information relating to them can lead to positive social change that fosters inclusivity and acceptance (Allport, 1954). Unsurprisingly, this approach was found to have the highest hashtag usage (10/10), as most external information is related to trending topics that are more closely aligned with news cycles (Zappavigna, 2011). As a result, the hashtags used in this context differ from those used in organisational contexts, as they pertain to trending topics being widely discussed. Non-organisational information can be relevant in local, regional, national, and international contexts. Examples are provided below:

*"Highest-paid #femaleathlete in the world , #NaomiOsaka is standing up for #Blacklives as she announced that she's pulling out of her #semifinalstennismatch today. 🗳️#BlackLivesMatter"--- **(Women Connect First)***

*"Well done to everyone who has shown up in person or joined in the twitter storm to oppose those seeking to stir up racism and hatred in Penally #LovePenallyHateRacism #RefugeesWelcome #Noplaceforhate" --- **(EYST)***

*"We're saddened to hear of the death of Senedd member Mohammad Asghar - the first ever BAME member of the Senedd and well liked. We send our condolences to his family." --- **(Diverse Cymru)***

*"Today brings mixed emotions for trans people and allies. Changes achieved together may not go as far as we'd hope, but are a small step forward despite increasingly regressive threats to trans rights. You've protected existing rights and edged us closer to trans equality." --- **(Stonewall Cymru)***

*"Happy Hanukkah to Jewish trans, nonbinary and intersex people and their allies!" --- **(Trans Aid Cymru)***

#### 6.4.1.4. Link to Other Content (non-social media)

Link to other content (non-social media) tweets represent 21.9% (n=626) of all push strategy tweets and 12.5% of all observed tweets in the study. This sub-strategy involves organisations sharing news articles, information, or fact sheets that are externally written to the organisation on other websites, often including a summary authored by the organisation. Like the non-organisation information sub-strategy, this approach relates to information that is not specific to the organisation and exists in various contexts. These can include news articles, important statistics, or other organisation's websites. In many cases links to news outlets are used to highlight current affairs, that the organisations condemn or support. The range of issues includes both domestic concerns, such as structural racism within public agencies, as well as international matters, such as restrictive policies imposed on LGBT+ communities in Poland.

*Health minister Vaughan Gething claims Welsh NHS has problem with 'structural racism' | Wales | ITV News <https://t.co/iN4Sqbd2T2>---(EYST)*



Figure 29. News article provided in Twitter link (1)

Figure 29:

*"It is condemning an entire community to hate." We need to continue to push for equality. As Poland is taking some shocking steps in the wrong direction, we need to remember that we're not equal until we're all equal*

👉 <https://t.co/D2zhT47ucq> #LGBTQ #pride #HumanRights"---  
**(Pride Cymru)**



Figure 30. News article provided in Twitter link (2)

In other cases, links lead to content that celebrates minority groups. This aims to positively influence audience perceptions by increasing exposure to positive content and countering negative biases. This finding is consistent with Allport's (1954) positive contact theory, highlighting the importance of intergroup interactions and exposure in promoting cohesion. For example, one post included a link to the 'pinc list' that celebrates influential members of the LGBT+ community:

"🌟 The #pinclist2020 is out!!! Meet the activists, politicians, students, artists, sport stars, and generally awesome people who are taking a stand for #LGBTQ equality across Wales - thank you for making a

difference. #pride @WalesOnline @JoeAli\_ <https://t.co/71qBJOyQIt>  
"--- (Pride Cymru)



Figure 31. Figure 31: News article provided in Twitter link (3)

In many ways this approach reflects the observational pull-strategy, however it is categorised as push because hyperlinks lead to other organisation's content. It is therefore unsurprising that this strategy ranked the highest (10/10) for hyperlink use. These links often lead to more comprehensive information than on Twitter due to its textual character restrictions (Han et al., 2019). This approach therefore pertains less to information reach, as seen in other push strategies, and more to information depth, allowing audiences to consume more in-depth and comprehensive information (Van De Velde et al., 2015).

#### 6.4.1.5. Link to Other Content (social media)

Link to other content (social media) was the most prevalent sub-strategy within push, accounting for 37.9% (n=1083) of push strategy tweets and 21.7% of all tweets in the study. This strategy shares similarities with the link to other content (non-social media) sub-strategy in terms of format and the type of information being shared,

often with a summary provided by the organisation. However, instead of providing a link to a news article on a different website, this strategy involves interacting with other tweets on Twitter, either by quoting or replying to them while composing a push tweet. The high prevalence of this strategy suggests a shift in the implementation of one-way push strategies to incorporate some dialogical elements, allowing organisations to network with other Twitter accounts while still broadcasting information (Greenberg and MacAulay, 2009). The extent of this is demonstrated by this strategy ranking second highest (9/10) for mentions among all strategies included in the study. This hybrid approach allows for disseminating inclusive content while strengthening ties and broadening relational space with other stakeholders (Kagan and Duggan, 2011). Consequently, the content tends to focus on more localised topics than those found in link to other content (non-social media). This can be attributed to organisations seeking to engage with stakeholders within their immediate purview who have the potential to influence their daily operations:

*"If you are an adult living in Wales and have been referred to the Gender Identity Clinic in London and haven't had an appointment, the Welsh Gender Service in Cardiff wants to contact you. Give consent to be contacted by using this form @UmbrellaCymru : <https://t.co/9hTKdY8hjk>" --- (Stonewall Cymru)*

*"Pride Cymru are working with Disability Sport Wales to develop a programme aimed at supporting disabled and non-disabled LGBTQI+ individuals into (or back into) physical activity across North Wales. Could you help by filling out their questionnaire? <https://t.co/uQaPXtOY1w>" ---(Disability Wales)*

*"This looks amazing!! Thank you @tfwrail ❤️ <https://t.co/Ckrc2Wn8x6>" --- (Pride)*

These demonstrate how organisations utilise broadcasting techniques to showcase positive outcomes, available services, or mere existence of other stakeholders. This practice broadens audiences' support networks and strengthens relationships with external stakeholders (Jue et al., 2010). This may explain the high ranking of emoji usage (7/10) in organisational tweets, as emotive emojis such as hearts are

frequently used by organisations to demonstrate positive interactions with other organisations.

#### 6.4.1.6. Pull Strategy

Pull strategy approaches align more closely with the dialogical than broadcasting paradigm (Greenberg and MacAulay, 2009). In this context, they are more designed to strengthen social capital with audiences (Granovetter, 1973). Mergel (2012) contends that organisations do not typically employ pull strategies to stimulate discussions on social media platforms. Instead, they are generally utilised to redirect audiences towards other forms of engagement outside the platform. These alternative forms of engagement still lead to strong social ties between audiences and organisations but are not limited to digital spaces. Pull strategy tweets promote engagement and participation, with organisations using followers and stakeholders as resources or information sensors (Mergel, 2012). The study found that pull strategy tweets accounted for 28.8% (n=1442) of tweets collected. Thus, pull strategy was the second most prevalent strategy in Mergel's (2013) typology, only behind push. Much like for push, pull strategy approaches were further explored in three sub-categories: (1) interactive, (2) work recruitment, and (3) observational.

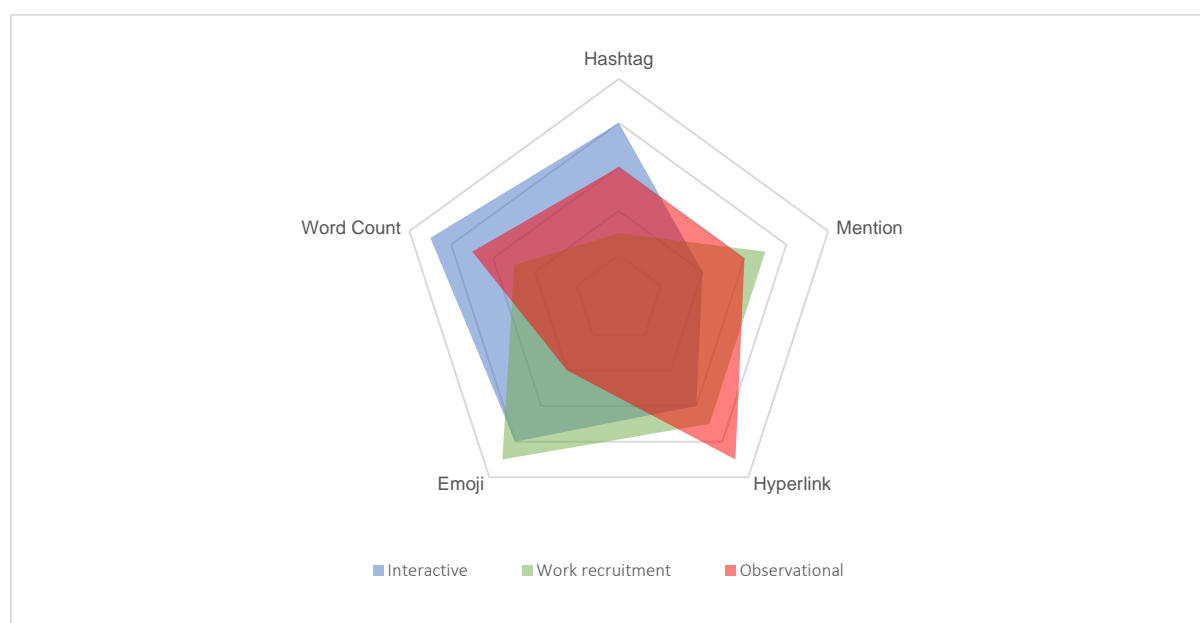


Figure 32: Pull strategy attribute radar graph

6.4.1.7. *Interactive Pull*

interactive tweets constituted the most prevalent form of pull strategy tweets, accounting for 56.9% (n=821) and 16.4% of all tweets analysed. Interactive pull strategy tweets were operationally defined as tweets that (1) encouraged active engagement with the organisation and (2) did not solely involve discussions on social media - instances, where social media was used for networking, were categorised as networking strategies. Subsequently, this approach consistently reflected the organisational information strategy by solely relating to organisation-specific content. Therefore, it was not surprising to find that this approach was ranked highly (8/10) for hashtag usage, which were also predominantly organisation-specific (e.g., #YourPride). These tweets are implemented in several ways, such as inviting followers to participate in activities, focus groups, town hall meetings, or to complete surveys designed by the organisation. While some of these tweets included hyperlinks to digital surveys, many others did not require any hyperlinks and instead featured images that provided details about the events, including locations and start times. Exemplars are presented below:

*"HAVE YOUR SAY Today: We want to hear from you, your views and concerns about COVID-19 will dictate our responses 🗳️ to Welsh and UK Governments, and public sector organisations. Survey here <https://t.co/IVs13DFYqA> <https://t.co/GPLSRJyDBg>" ---*  
**Diverse Cymru**

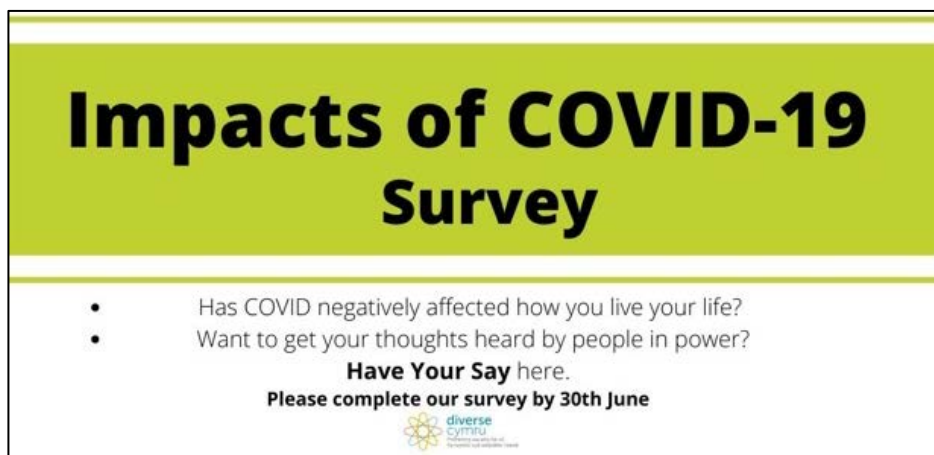


Figure 33. Tweet image with event information (1)

"Come talk to us about how we can make our services more accessible to your community Dewch i siarad â ni am sut y gallwn ni wneud ein gwasanaethau'n fwy hygyrch i'ch cymuned 10th Feb, 2pm-3pm <https://t.co/jilVC0vDhH> 10th Feb, 5:30pm - 6:30pm <https://t.co/gp5Zzkh13H> <https://t.co/yRriSu57e4>" --- **WWA**



Figure 34: Tweet image with event information (2)

"There's just under a week until our first event of the year! The Big Queer Quiz is funded by Comic Relief's LGBTQ COVID-19 Recovery Fund, in partnership with METRO Charity and Umbrella Cymru, and will take place next Saturday at 12pm! Get your tickets here <https://t.co/ns6xezPfzv> <https://t.co/q2c07CNX4V>" --- **Trans Aid Cymru**





Figure 35. Tweet image with event information (3)

*"Start your Sunday the right way with some yoga at 10:30! Head over to our Facebook page to join @somethingtosi! #YourPride <https://t.co/HPkmyjOuDH>" --- **Pride Cymru***

These show interactive approaches can serve two distinct purposes. While both objectives involve active participation, the first two align more closely with traditional pull strategy techniques, as Mergel (2012) outlined, utilising audiences and communities as information sensors to gain valuable insights that inform organisational operations. This is often achieved through public consultations or by providing surveys. The latter two excerpts illustrate that interactive approaches can also be utilised to enhance social capital with audiences (Granovetter, 1973) by fostering a sense of community and organising events that serve purposes beyond gathering information, such as hosting quizzes or yoga sessions. Alongside ranking highly for hashtags, this approach also ranked high in word count (9/10) this can often be seen, with tweets being presented in longer format to fully detail the interactive activity being promoted (i.e., start time, location, and further information).

#### 6.4.1.8. Work Recruitment Pull

Work recruitment tweets accounted for 11.9% (n=171) of pull strategy tweets and 3.4% of all tweets, positioning it as the least frequently used form of pull strategy tweets. However, despite its lower ranking, work recruitment tweets are still more prevalent than any transactional approach. These findings align with recent evidence highlighting the growing significance of social media platforms as a conducive environment for promoting paid or voluntary work opportunities (Olinski and Szamrowski, 2021). Work recruitment tweets were based on explicit calls to action encouraging readers to apply for paid or volunteering positions within the account's organisation. It is worth noting that tweets promoting paid or volunteering positions with other organisations were not included in the work recruitment category. Instead, they were classified as 'link to other content (non-social media).

*"For more information/ to apply for the role please find contact details above 🙌🙌 #Deliverydriver"--- **Women Connect First***

*"Check out our amazing job opportunities on @CharityJob Work with wonderful people while doing amazing things!  
<https://t.co/ghAZccLiu4> <https://t.co/T8oZ3VWREk>"--- **Pride Cymru***

*"3 Days to Go !!!!! Are you passionate about social media and communications ? We are looking for an enthusiastic individual to join our team as a Social Media Coordinator @ 15hrs per week.  
@charityjobfindr @cardiff\_jobs @cdfvolcentre @DC\_Wales  
@ProMoCymru @BMESkillsWales <https://t.co/oFWW11YMHd>"---  
**Women Connect First***

Recruitment tweets rank high for emojis (9/10), mentions (7/10), and hyperlinks (7/10). Upon inspection of tweet extracts, symbolic emojis such as arrows and pointing fingers are used to direct audiences to application links, encouraging and facilitating application submissions. This finding supports Highfield's (2018) earlier research, which suggests that emojis can be used to increase engagement in brands, events, and promotional campaigns. It further extends this notion to include their potential for encouraging applications for roles. Although the high ranking of mentions may initially seem confusing, closer examination of the tweet extracts

reveals that organisations mention job pages (e.g., @cardiff\_jobs) that they also use to advertise the roles on. A high ranking in hyperlinks is expected, and a quick inspection of the qualitative output reveals that hyperlinks are used either to direct readers to more information about the role or to direct them to the actual application.

#### 6.4.1.9. Observational Pull

Observational tweets constitute 23.9% (n=345) of pull strategy tweets and 6.9% of all tweets in the study. They redirect followers to external sources of information, typically hosted on the organisation's website. Unlike interactive strategy tweets, observational tweets do not encourage active participation but encourage the consumption of information produced by the organisation. Examples of the content found in observational tweets include information sheets, new legislation, newsletters, and details about the work carried out by the organisation. Selected examples of observational content are presented below:

*"This #InternationalWomensDay, we want to highlight the work of lesbian, bi & trans women activists worldwide. Discover the new website for our #OutOfTheMargins project, which empowers local organisations to improve their lives around the world. #IWD  
<https://t.co/mTlnrEo9AI> <https://t.co/z4edploA12>"--- **Stonewall***

*"Our monthly newsletter is out now!! Read all about what we got up to supporting BAME people in July!! Featuring @childcomwales @HelenaHerklots @CardiffUniInter @CABCymru @WelshGovernment @SwanseaCouncil @childrensociety and more!! <https://t.co/sTDTOr4Rkl>  
#WeareWales 🍷 <https://t.co/y1pTn4adCn>"--- **EYST***

*"Many Muslims are preparing for Ramadan - we've compiled some information and resources in support. Ramadan Mubarak. 🇬🇪 <https://t.co/hLwfICMZUZ> <https://t.co/NvEVJbmePj>"---  
**Diverse Cymru***

*"Given the challenges of 2020, it feels right to reflect on the amazing work of 2019. Read our Annual Report here <https://t.co/vbOEKEggG5> <https://t.co/ZmwCWsVveA>"--- **Welsh Refugee Council***

While most examples direct audiences to the organisation's website, some lead to external platforms that host the organisation's content, such as YouTube, which utilises audio-visual modalities (Waterloo et al., 2018). Hence, observational pull strategies not only promote reading, but also viewing content.

*"Will EU Citizens still be able to access healthcare in the UK after Brexit? If you are a EU Citizen living in Wales it's important you apply to the EUSS scheme. Learn more here: <https://t.co/n2d07h5Jft> #brexitwhatnext #EUCitizens"--- **Disability Wales***

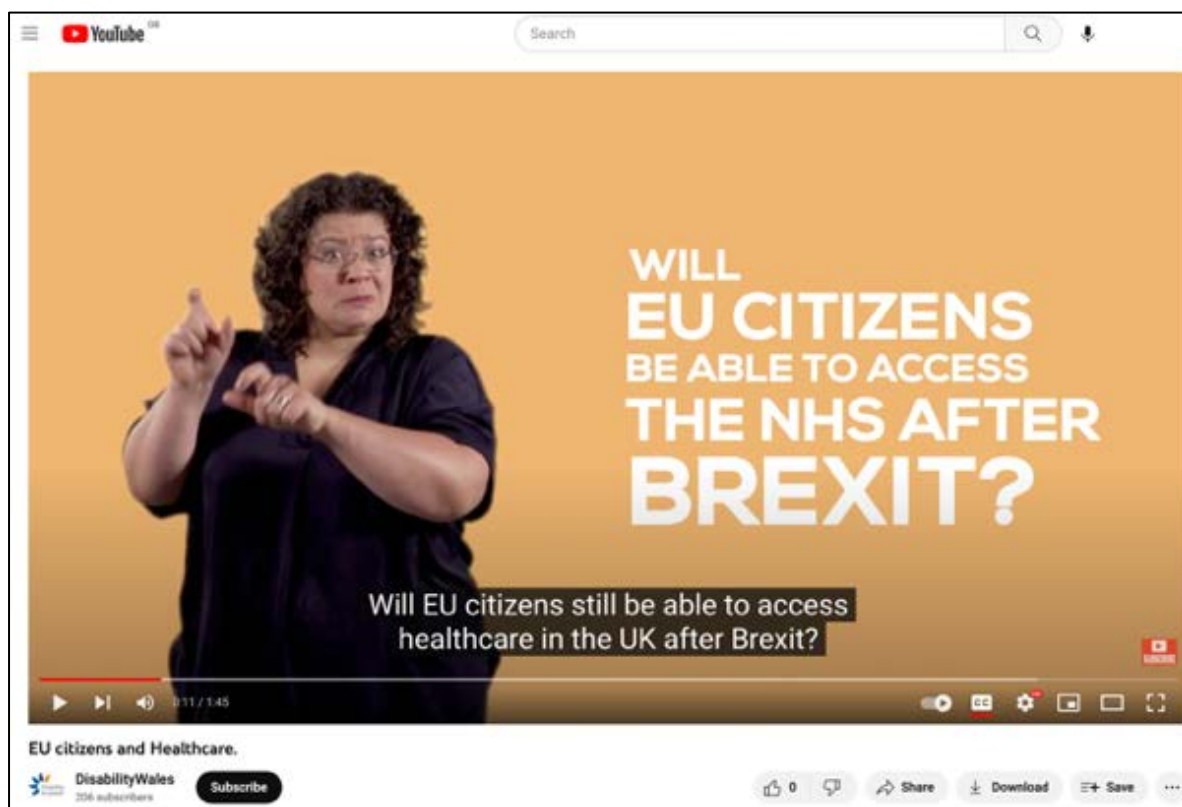


Figure 36: Link to organisation's YouTube content

Subsequently, observational approaches exhibit less dialogical features than any other pull strategy, and prioritise information diffusion, similar to one-way

broadcasting approaches (Greenberg and MacAulay, 2009). Similar to link to other content (non-social media) tweets, this strategy facilitates information depth (Van De Velde et al., 2015) by redirecting audiences to other websites that do not have character restrictions (Han et al., 2019). As such, it is unsurprising that this approach ranked second highest (9/10) for hyperlink usage, behind only link to other content (non- social media).

#### *6.4.1.10. Networking Strategy*

Networking tweets accounted for 13.0% (n=650) of all tweets. In Mergel's (2013) typology, networking approaches move beyond initial two-way pull dialogical approaches towards an even more interactive, two-way cross-boundary framework. Such interactions enhance social capital and enable a round-the-clock interactive environment where organisations can strengthen existing relationships with constituents and other stakeholders while establishing new ones. Lovejoy and Saxton (2012) highlight the potential of social media in facilitating community building and the establishment of 'issue networks'. Organisations network with public audiences, resulting in an interactive co-design of services. Social media has enabled access to instant high-quality content and global networking opportunities (Berlot et al., 2012). Networking strategy tweets were identified as any tweet directly promoting dialogical engagement on Twitter. Content that promotes engagement on other agency-run websites is coded into one of the pull sub-strategies. Networking strategy tweets predominantly reply to other users on the platform, ask general questions, and use polls to instigate online discussions. They can be addressed to individuals, groups, or all followers. For example, a networking strategy tweet may thank users for their support and encourage them to spread the word or ask followers about their opinions on a current issue relevant to the organisation's mission.

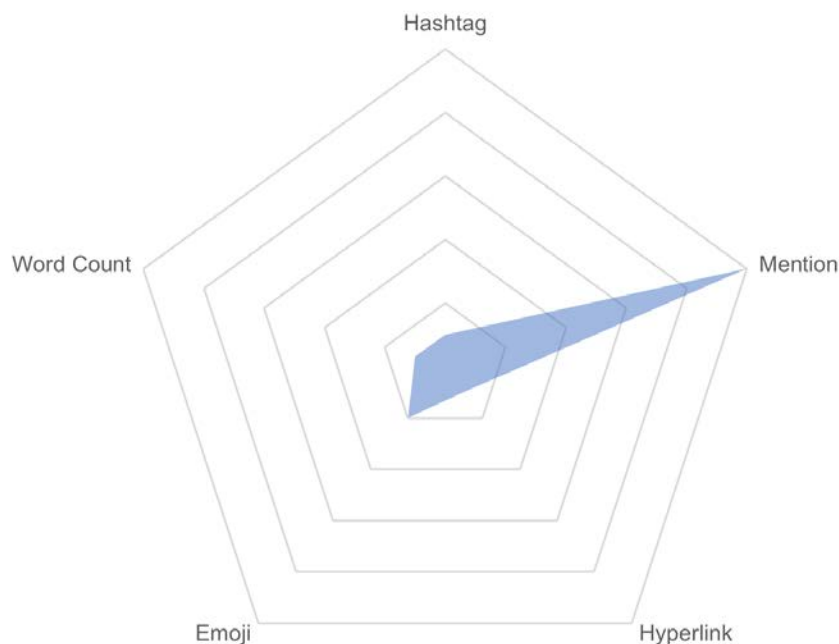


Figure 37: Networking strategy attribute radar graph

Networking approaches are characterised by certain sensory elements of the pull strategy, as they involve seeking input from audiences on the operations of organisations. However, the key distinction lies in the additional interactive dynamic afforded by the instantaneous communication capabilities of social media platforms. Examples are given below:

*“As part of 2020’s digital pride, we want to create #YourPride playlist. What are the anthems that you blast as you’re getting ready for the festival? Tweet us your suggestions below and we’ll reveal the playlist later in the month.” --- **Pride Cymru***

*“@Becky18324396 Hi Becky. We did not ignore you. We told you that we would consider putting that option in future surveys. We are sorry that you didn’t feel able to fill out our form”. --- **Trans Aid Cymru***

*“@ACCSwansea Thank you” --- **EYST***

*“@cavamh Happy birthday from us! ^Joe” --- **Diverse Cymru***

Thematic analysis found that networking communications in the third sector share similar characteristics with those identified in the literature. Meijer and Thaens (2013) describe networking communications as decentralised, indicating that their objectives and tasks are often unique and diverse. This approach reflects the co-design of services, which enables organisation constituents to have an equal say in the interactions, and independent queries made by individuals to be addressed (Mergel, 2013). Unlike push and pull strategies, which are characterised by tone-setting dynamics, networking communications do not involve broadcasting or pulling audiences towards broad concepts but instead address individual queries. Consequently, networking strategies are not further categorised into sub-strategies. The highly interactive nature of networking approaches strongly emphasises mentions as a communication attribute. This is reflected in their ranking, as mention use ranks highest of any strategy (10/10), while networking ranks lowest in the remaining attributes. Notably, the word count associated with networking is also lower than other strategies. This is evidenced by extracts that depict long conversations occurring over multiple posts within a single thread.

#### *6.4.1.11. Transactional Strategy*

The final strategy in Mergel's (2012) typology is the 'transactional'. This approach involves organisations utilising online communication to transact with external stakeholders or audiences regarding physical goods or services (Mergel, 2012). While transactional approaches are inherent in the private sector (Thomas, 2007), they are only recently emerging in the public sector. In the public sector, transactional approaches relate more to providing public services than goods (Mergel, 2012). In the third sector, initial evidence indicates that transactional approaches are primarily used to elicit donations, with varying levels depending on the type of organisation. For instance, disaster relief non-profits tend to appeal at a higher rate than educational organisations that mainly broadcast informational content on policy (Deschamps and McNutt, 2014). As suggested by Mergel (2013), the prevalence of transactional strategies is lower than that of the other three broad strategies, accounting for only 3% (n=153) of tweets observed in the study.

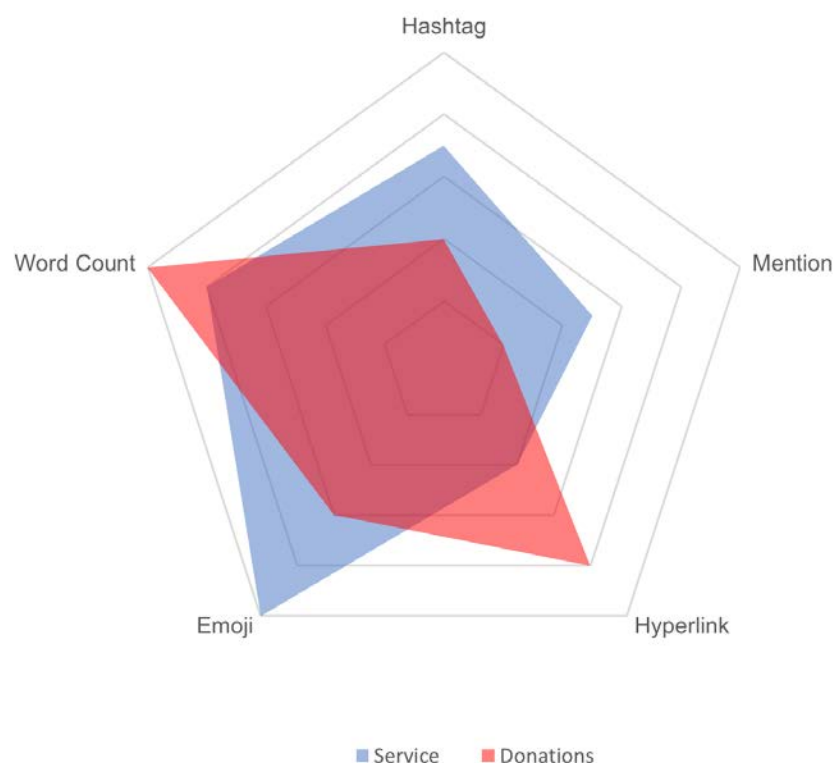


Figure 38. Transactional strategy attribute radar graph

#### 6.4.1.12. Elicit Donations

The first sub-strategy involves organisations seeking to solicit donations on social media platforms. These requests were made through simple appeals for donations or through incentives, such as prize giveaways, trivia contests, and product sales, as demonstrated in the final extract. This finding aligns with existing literature that suggests non-profits are exploring innovative ways to raise funds (Castillo et al., 2014). Nevertheless, soliciting donations accounted for the lowest frequency in the data (n=42), significantly less than other non-profits that primarily rely on donations, such as disaster relief (Deschamps and McNutt, 2014). Some posts explained how donations would be used, while others included motivational factors, such as the importance of being an 'ally' or framing the cost of donations relative to other expenditures. Using such approaches to persuade audiences may explain why the word count for donation-related posts was the highest of any strategy (10/10). These posts also received relatively high scores for hyperlinks (8/10), which redirected



readers to other web pages providing more information on how donations would be utilised and how to donate. Examples of such posts are provided below.

*“Allies can support Welsh trans & non-binary people with financial needs by donating below & sharing: <https://t.co/c55c91rgQl> “ ---*  
**Trans Aid Cymru**

*“It’s £23.75 a week or just £3.39 a day. That’s less than the cost of a Meal Deal or most takeaway coffees. Your support can help us support those in hardship. Donate here: [https://t.co/sbGJ6McKbM\\_](https://t.co/sbGJ6McKbM_) -*  
**-- Welsh Refuge Council**

*“We’re working hard to support vulnerable BME Communities in Wales, and to represent their views, experiences and opinions at a strategic level. Donate to support our work here -*  
<https://t.co/XhLPsGxOiD> <https://t.co/rXdZPKZero> “ ---  
**Diverse Cymru**

*“Our fabulous patron Dr Bev has released this much-anticipated line of merch for you to finally get your hands on! £1 from each item sold in the Enfys range will be donated to Pride Cymru. Which items will you be snapping up first?*  
*@Keetch1701 <https://t.co/RqLbJuK9Jn> [https://t.co/MHRcgASAMi\\_](https://t.co/MHRcgASAMi_)*  
**--- Pride Cymru**

#### 6.4.1.13. Services

The other transactional strategy is ‘service use’. Service tweets were more prevalent than (n=105) of tweets eliciting donations. This was a surprise as most existing literature indicates that donations are the primary form of NGOs’ transactional approaches. As such, this finding suggests that pro-minority rights third sector groups utilise transactional approaches more akin to the public sector, in which they promote services to audiences. Services in this context entirely related to support services such as helplines provided by organisations. These are either provided by hyperlinks to support service pages on organisation websites, or by simply providing the hotline number. Hyperlinks are therefore less integral to this sub-strategy as

demonstrated in the ranking (4/10). Support services varied between agencies, but broadly relates to, general advice, victim support, legislative support, and informational services. Examples are given below:

*“Diverse Cymru is offering: ✓ Free ✓ Award-winning ✓ Independent support throughout Cardiff and the Vale. Get help with benefits, PIP, ESA, and challenging decisions. Contact us at [advocacy@diverse.cymru](mailto:advocacy@diverse.cymru) today. <https://t.co/qJHU4Cl2Sb>”---*  
**Diverse Cymru**

*“#BAMEHelpline 📞 Never be afraid to ask for help Here's a response from just one of our clients - and we were so thrilled that we could offer the help they needed! We're open Mon-Fri 10:30am-2:30pm - will you help us share? <https://t.co/X14Ofzrag>” ---*  
**EYST**

*“Our Advice & Advocacy Service for our Golden Years Project is available over the phone We are always here and will continue to help and support you during this crisis period @mikelewis1960 @LBFEW #StayHomeStaySafe #Togetherwecan <https://t.co/4g8iidEk8h>” ---*  
**Women Connect First**

*“@Welshrefcouncil is open for business. Call us on: 02920 489 800 for Cardiff 01633 266 420 for Newport 07918 403 666 for Swansea 07977 234 198 for Wrexham 📞 📞 📞 We have caseworkers ready to help #SanctuaryandSafety #StayHome #StaySafe <https://t.co/ekxcL6hSAi> <https://t.co/hfmhUjL8Kc>”---*  
**Welsh Refugee Council**

Services ranked the highest of any strategy (10/ 10) for emoji use. A closer inspection of examples indicates that this outcome can be largely attributed to the use of phone emojis, commonly positioned near phone numbers. This strategic placement is likely aimed at drawing attention to the phone number, enhancing the probability of followers availing the services.

Table 18. Negative binomial regression, predicting digital engagements

	Likes				Retweets				Replies			
	B	SE	Wald	Exp(B)	B	SE	Wald	Exp(B)	B	SE	Wald	Exp(B)
Ref: Push												
Pull	-0.37	0.04	96.69	0.69***	0.07	0.04	3.20	1.07*	-0.18	0.08	5.52	0.83***
Networking	-1.03	0.07	240.51	0.36***	-1.46	0.09	290.86	0.23**	0.04	0.12	3.10	1.04**
Transactional	-1.08	0.17	41.98	0.34***	-0.85	0.18	21.60	0.43**	-1.07	0.39	7.43	0.34***
Word Count	0.02	0.00	315.67	1.03***	0.03	0.00	493.16	1.04***	0.03	0.00	137.96	1.03***
Hashtag No.	-0.10	0.02	25.28	0.91***	-0.13	0.02	40.29	0.87***	-0.15	0.05	10.02	0.86***
Mention No.	0.10	0.01	97.73	1.11***	0.10	0.01	82.13	1.11***	0.03	0.02	3.03	1.04*
Hyperlink	0.50	0.05	99.19	1.65***	0.80	0.06	194.38	2.23***	-0.27	0.09	8.27	0.77***
Mention Covid	-0.20	0.05	13.48	0.82***	-0.02	0.06	0.18	0.98	-0.18	0.11	2.57	0.84*
Celebration	0.28	0.08	11.77	1.32***	0.14	0.09	2.55	1.16	0.18	0.16	1.19	1.19
Intervention	0.84	0.06	138.28	2.30***	0.67	0.14	62.38	1.95***	0.05	0.09	18.28	1.05**
Other Focus	-0.06	0.06	0.78	0.94	0.03	0.07	0.19	1.03	-0.10	0.14	0.52	0.91
Ref: High Tweet Frequency												
Low	-0.69	0.05	199.53	0.50***	-0.81	0.05	216.11	0.45***	-0.58	0.11	28.52	0.56***
Mid	0.66	0.06	106.79	1.94***	0.09	0.07	1.97	1.10	0.96	0.14	49.92	2.61***
Ref X-Large Followers												
Small	0.96	0.07	208.70	2.60***	0.83	0.07	141.97	2.29***	1.18	0.14	67.09	3.25***
Medium	0.39	0.07	35.26	1.47***	-0.09	0.07	1.71	0.91	0.52	0.14	13.30	1.68***
Large	1.13	0.08	193.07	3.10***	0.64	0.09	51.48	1.89***	0.77	0.18	17.20	2.15***
Ref: No Emoji												
Activity Emoji	0.20	0.09	5.68	1.22***	0.20	0.09	4.75	1.22***	-0.05	0.19	0.07	0.95
Symbol Emoji	0.01	0.09	0.02	1.01***	0.16	0.09	3.12	1.18***	-0.10	0.20	0.24	0.91
Emotive Emoji	0.59	0.10	34.55	1.81	0.68	0.11	37.36	1.98*	0.17	0.18	0.89	1.19
Group Specific Emoji	0.45	0.11	17.92	1.57**	0.44	0.12	14.35	1.55**	0.29	0.20	2.05	1.34
Knowledge #	0.18	0.06	10.82	1.20***	0.37	0.06	39.36	1.45***	0.08	0.12	0.41	1.08
Event #	0.19	0.07	7.74	1.21***	0.13	0.08	3.21	1.14*	0.21	0.15	1.99	1.23
Time/ Place #	0.12	0.09	1.69	1.13	0.30	0.10	9.23	1.35***	0.46	0.18	6.16	1.58***
Branding #	0.05	0.07	0.51	1.05	-0.14	0.08	3.20	0.87*	0.09	0.15	0.34	1.09**
Call-to-Action #	-0.20	0.11	3.31	0.82*	-0.11	0.12	0.96	0.89	-0.70	0.29	5.92	0.50
Dialogical #	0.09	0.12	0.40	1.07	-0.22	0.12	2.54	0.78	0.67	0.14	37.28	1.99***
Goals #	0.19	0.07	7.44	1.21***	0.08	0.08	1.06	1.08	-0.04	0.16	0.05	0.96
Weekend	-0.02	0.05	0.12	0.98	-0.05	0.05	1.07	0.95	0.03	0.09	0.12	1.03
Ref: Evening												
Night	0.04	0.12	0.14	1.04	-0.06	0.13	0.17	0.95	-0.30	0.23	1.71	0.74
Morning	-0.01	0.05	0.07	0.99	0.15	0.06	7.00	1.16***	-0.22	0.10	4.53	0.80**
Daytime	-0.04	0.04	0.78	0.96	0.09	0.05	3.60	1.09**	-0.19	0.08	5.39	0.83**
N	N= 5000				N= 5000				N= 5000			
-2 Log Likelihood	-13942.26				-10936.13				-3051.83			
Df	32				32				32			

### ***6.4.2. Digital engagement models***

Most social media research has focussed solely on retweets when exploring engagement data<sup>62</sup>. This study provides an overview of third sector minority rights organisation's use of social media as a whole. It therefore measures discrepancies in a wider range of engagement types (i.e., likes and replies). This marks the first study to measure Mergel's (2013) communication framework in relation to engagement data. The literature suggests that on-platform engagement types can have significantly different meanings. For instance, retweets are typically linked with endorsement and the degree to which digital audiences want to propagate and amplify particular ideas posted on social media (Williams and Burnap, 2016).

Conversely, likes are less designed to amplify opinions but rather to simply just endorse them. In this context, likes can be regarded as a more passive form of endorsement than retweets. While likes and retweets share some similarities, replies differ significantly from them. Replies are primarily associated with on-platform open discourse and can reflect social capital, highlighting the level of verbal input and engagement within specific digital communities (Burton and Soboleva, 2011). Notably, replies are not always indicative of positive social capital, as they may also be used to challenge initial posts, as demonstrated in Chapter Five. Three models were developed in this chapter, each with the same independent variables and different dependent variable relating to the three on-platform engagement metrics.

#### *6.4.2.1. Sub Model Analyses*

Sub-model analyses were conducted for all three models. Instead of using statistical procedures such as PCA that consider correlations, the study manually identified five factor groups: content factors, account factors, hashtag and emojis, controls, and strategy factors. The detailed sub-factor analysis results are provided in Appendix 10

---

<sup>62</sup> The three primary on-platform engagement metrics on Twitter are retweets, likes and replies (Boyd et al., 2010)

which demonstrate variations in the degree to which the factor groups account for the variance in each model. Content factors had the greatest influence on likes<sup>63</sup> and retweets<sup>64</sup>, followed by strategy factors, account factors, hashtag and emojis, and controls, respectively. The results suggest that on-platform modalities, such as hyperlinks, mentions, and word count, are instrumental in driving engagement on Twitter. Content that includes hyperlinks or mentions can also increase engagement by facilitating the discovery of related content and encouraging interactions with other users. This suggests that NGOs can leverage these factors to promote community cohesion on social media.

The study also found that strategy factors, which reflect Mergel's communication framework of push, pull networking, and transactional, were the second most influential factor group for likes and retweets. These findings indicate that strategies can significantly influence propagation and endorsement of positive and inclusive content. However, their impact is not as significant as on-platform content modalities. Account factors, such as tweet frequency and account size by followers, were the third most influential factor group for likes and retweets. This suggests that users may be more likely to engage with content from accounts they perceive as active and influential. However, such factors play a significantly lesser role than strategic approaches or content modalities. Emojis and hashtags had minimal impact on likes and retweets, indicating that their use may not be as crucial for driving user propagation and endorsement as previously thought. Controls, such as time of the day or day of the week, also had a limited impact on engagement for likes and retweets, suggesting that the timing of posting may not be as critical as the quality of the content, overall strategy, or account. The results of the sub-model analysis for

---

<sup>63</sup> Content factors explained the most variance (-2 Log-Likelihood = **-14383.32**, BIC= **28843.29**), followed by strategy factors (-2 Log-Likelihood = **-14546.01**, BIC= **29126.09**), account factors (-2 Log-Likelihood = **-14580.22**, BIC= **29211.55**), hashtag/ emoji type (-2 Log-Likelihood = **-14741.65**, BIC= **29585.50**) and finally the controls (-2 Log-Likelihood = **-14806.37**, BIC= **29655.31**).

<sup>64</sup> Content factors explained the most variance (-2 Log-Likelihood = **-11419.49**, BIC= **22915.63**) followed by strategy factors (-2 Log-Likelihood = **-11734.72**, BIC= **23503.51**), account factors (-2 Log-Likelihood = **-11914.83**, BIC= **23880.76**), hashtag/ emoji type (-2 Log-Likelihood = **-12133.04**, BIC= **24385.31**) and finally controls (-2 Log-Likelihood = **-12193.91**, BIC= **24430.41**).

the reply<sup>65</sup> model differed significantly from the likes and retweets models, with account factors being the second most influential factor group, followed by hashtag and emoji type, controls, and then strategy factors. As such, factors driving on-platform dialogue may differ from those impacting general endorsement-seeking behaviours. In particular, users may be more likely to engage in dialogue with accounts, they perceive as trustworthy or credible. Surprisingly, strategy factors explained little variance for replies, which contradicts associated research that dialogical approaches are extremely influential in garnering on-platform discourse. Results later in this section will explore discrepancies in strategies in relation to replies and demonstrate that perhaps not only dialogical approaches are effective in receiving on-platform replies.

#### 6.4.2.2. Broad Strategy Results

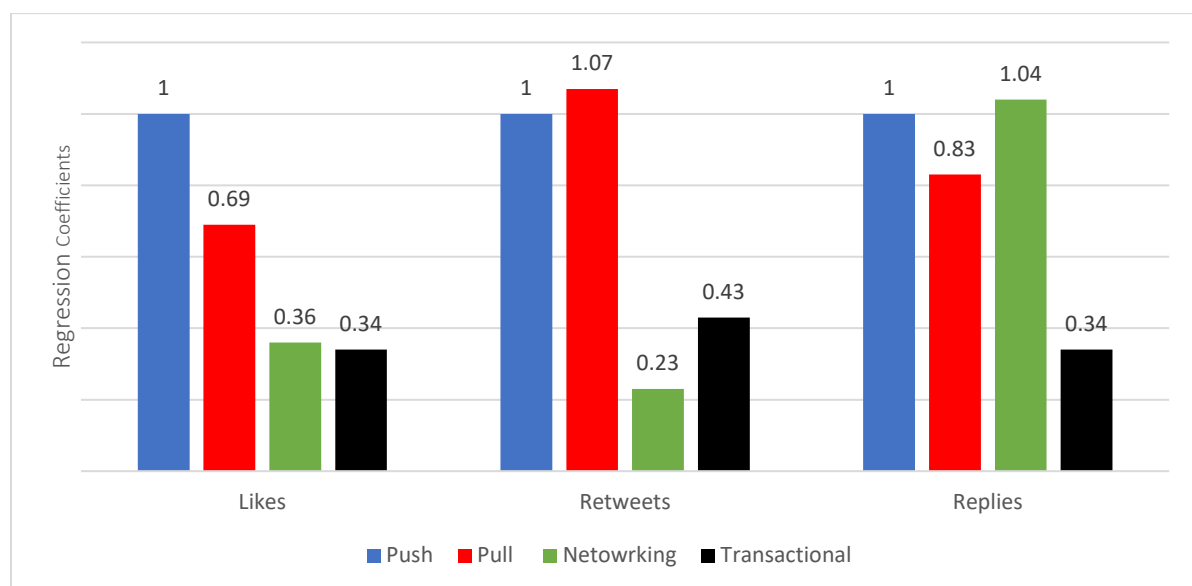


Figure 39. Clustered column chart showing regression coefficients for strategy factors (all three models)

<sup>65</sup> Content factors explained the most variance (-2 Log-Likelihood = **-3173.12**, BIC= **6422.88**) followed by account factors (-2 Log-Likelihood = **-3175.56**, BIC= **6402.22**), hashtag/emoji type (-2 Log-Likelihood = **-3266.50**, BIC= **6652.24**), controls (-2 Log-Likelihood = **-3276.46**, BIC= **6595.50**) and finally strategy factors (-2 Log-Likelihood = **-3295.87**, BIC= **6625.81**).

The clustered column chart in Figure 39 displays the strategy factor predictors for each model<sup>66</sup>. As noted by Mergel (2013), the strategies employed by organisations address varying key objectives and missions, and single organisations may employ multiple strategies at any given time. Upon inspecting the plots, it is evident that no single strategy is the highest predictor of engagement across two or more models, indicating that different strategies may be more beneficial depending on the desired outcome. Push strategy tweets dominate the likes model, with other strategies being relatively less associated. This result was expected, as push strategies are situated in the broadcasting paradigm, where one-way interactions are actively pursued, and information is broadcasted to audiences with little to no intention of building close-knit social ties. While broadcasting significant levels of information can promote positive content and expose wider audiences to other communities, allowing for better intergroup understanding, likes are most suitable for general information broadcasting as other engagements, such as replies, result in two-way dialogical interactions that serve different purposes (Greenberg and MacAulay, 2009).

Although push strategies remain effective in garnering retweets, pull strategies were more positively associated, perhaps due to their greater information depth and semi-interactive nature, resulting in a greater likelihood for audiences to propagate information posted by organisations (Van De Velde et al., 2015). Similarly, networking, and transactional strategies were significantly less associated with retweets than push and pull. However, transactional strategies have a higher likelihood of being retweeted than networking, where the opposite was true in the likes model. This finding implies that transactional practices such as appeals for donations or the provision of support services are more likely to be propagated by followers, allowing for a wider reach of such functions, which is crucial in both fundraising and ensuring that invaluable support services may reach a broader audience of needing individuals.

In the final model, transactional strategies were less likely to predict replies relative to other coefficients, and networking strategy was the highest predictor of replies

---

<sup>66</sup> The reference category in all three models is push strategy (due to its highest overall prevalence).

from all broad strategies. This finding is significant as it highlights the importance of networking strategy in spurring online discourse with the organisation's audience. Dialogical strategies that offer two-way communications to build strong relationships with audiences and associated stakeholders (Deschamps and McNutt, 2014) are critical in promoting community cohesion through social media as they help build a sense of community. Although networking strategy is closely followed by push strategy and then pull strategy, this result emphasises that networking strategy, which is based on the dialogic paradigm, is two-way interactive by design and actively seeks to cause open conversation between organisations and communities and external partners (Greenberg and MacAulay, 2009). These findings only provide initial and general understandings of social media strategies in enhancing community cohesion. Later in this chapter, further sub-strategies are explored that help shed light on more specific approaches taken by NGOs to promote community cohesion.

#### *6.4.2.3. Like Model*

Content factors exhibited both positive and negative associations. Controlling for other variables, interventions were found to have the most substantial positive impact, followed by hyperlinks, celebrations, mentions, and word count. Conversely, COVID-related tweets exhibited the most negative association with likes, trailed by tweets containing hashtags. Tweets referring to minority groups not typically focused on by the organisations did not show significant associations with likes, retweets, or replies across all three models. The account factor predictors showed significant associations with likes. Compared to the reference category of high tweet frequency, tweets with low tweet frequency were less likely to be liked, while tweets with medium tweet frequency were more likely to be liked. Moreover, tweets from accounts with small, medium, and large followers were more likely to be liked compared to the reference category of x-large accounts. The extent of this association varied, with large accounts being the most likely to receive likes, followed by small and medium accounts. Regarding the use of emojis and hashtags, tweets with group-specific emojis, activity, and symbolic and group-specific emoji use were more likely to be liked compared to the reference category of no emoji use. The likelihood of likes decreased in the given order. However, emotive emojis,



time/place, and branding hashtags did not show significant associations with likes. Informative, event, and goals/values hashtags were found to have positive associations with likes to similar extents, while organisation and call-to-action hashtags<sup>67</sup> showed negative associations with likes. Finally, the controls sub-factor showed no significant associations with likes, indicating that the day of the week or time of day did not take significantly affect.

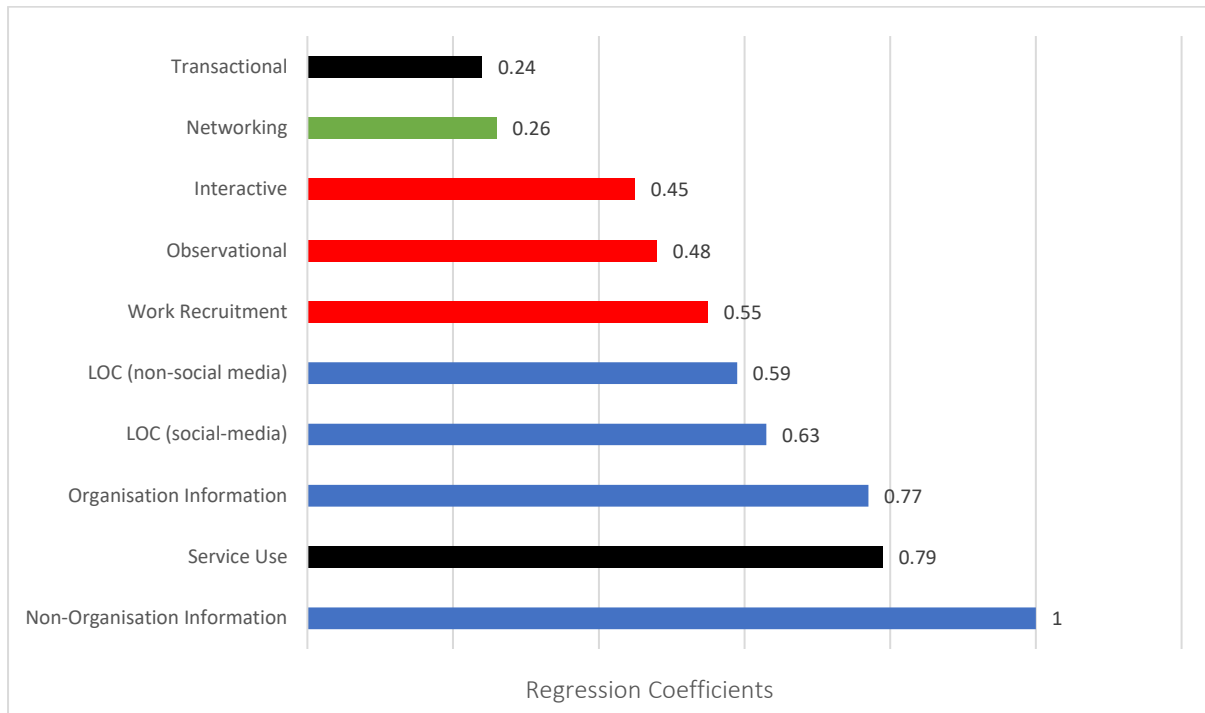


Figure 40: Audience engagement coefficients for 'like' sub-strategy model

Subsequent revised models were executed for all three dependent variables - likes, retweets, and replies - with the sub-strategies<sup>68</sup> included to provide more detailed insight into discrepancies in engagement. As previously mentioned in the broad strategies section, push strategy tweets were most associated with likes, followed by pull, networking, and transactional strategies. A quick examination of Figure 40

<sup>67</sup> Call-to-action (like model) hashtag predictor only approached conventional standards of statistical significance (\*)

<sup>68</sup> The selection of a reference category for the sub-strategies differed from the broad strategies approach, which relied on prevalence. Instead, the strategy (non-organisation information) with the highest mean engagement score was chosen to improve the cohesion of the results and test the potential effects of other factors included in the regression model.

reveals that push strategies dominate the higher end, indicating the strongest coefficients, thereby reinforcing the notion that push strategies primarily operate in a broadcasting paradigm where information is disseminated and consumed, rather than interacted with.

Interestingly, service use was the second highest sub-strategy, indicating that audiences widely accept support services organisations provide on social media. In contrast, donations are the least associated with likes. While this may suggest a limited endorsement of donation appeals, one possible explanation could be that endorsement may come in the form of making donations rather than liking posts. The remaining dialogic and semi-dialogic approaches belonging to pull and networking strategies are observed in the lower end of the chartle, suggesting that passive endorsement such as likes is less associated with them.

#### *6.4.2.4. Retweet Model*

Despite explaining the most variance, some content factors were found to have no significant associations, such as tweets that mention COVID-19, "other focus, or celebrations. Reflecting the likes model, the strongest positive association within content factors for the retweet model is observed by some margins in intervention tweets. This was followed by hyperlinks, mentions and word count. Interestingly, hashtags were the only content factor negatively associated with retweets, this reinforces the idea that pull strategy approaches focus more on information depth by redirecting audiences to other resources than information reach, by enhancing on platform discoverability (Van De Velde et al., 2015). As such, hyperlinks are effective tools in gaining retweets. This finding coincides with strategy factor results that show pull strategy tweets are more likely to be retweeted than push strategy tweets. Conversely, networking, and transactional were less associated with retweets than push strategy tweets. Concerning the tweet frequency reference category (high tweet frequency) low tweet frequency accounts were less likely to be retweeted. No significant association was found for mid tweet frequency accounts. In relation to the account size reference category (x-large accounts), small and large accounts were both more likely to be retweeted, with the association being greater for small than

large accounts. No significant association was found for medium-sized twitter accounts. In relation to the reference category 'no emoji use', all four emoji were found to have significant positive associations with retweets, with emotive having the most, followed by group-specific, activity and symbolic. From hashtag types, only goals/ values and call-to-action hashtags were found to have statistically insignificant associations. Positive associations exist between retweets and informative, event<sup>69</sup> and time/ place hashtag use. Conversely, negative associations exist for organisation and branding hashtag types, although in both cases these only approached conventional standards of statistical significance. From controls, days of the week had no significant association with retweets. In relation to the reference category of evenings, morning and daytimes were more likely publish times to receive retweets.

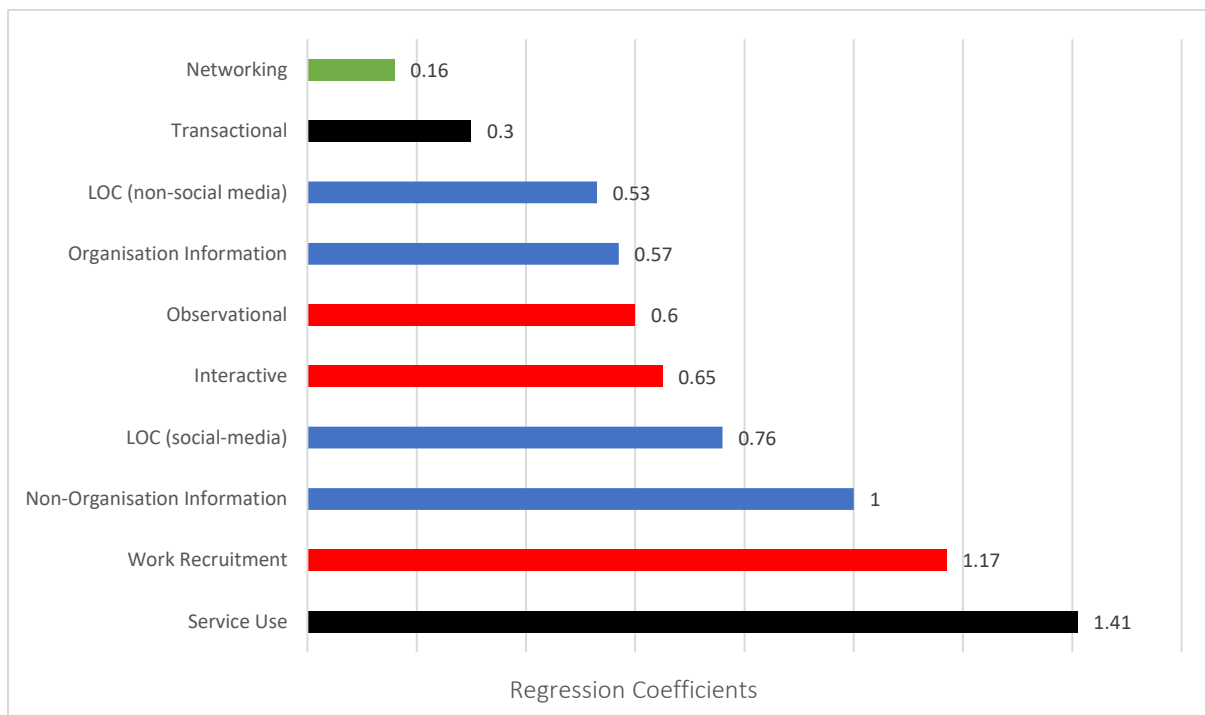


Figure 41: Audience engagement coefficients for 'retweet' sub-strategy model

Upon examination of the retweet coefficient rankings, a notable difference is observed with respect to the reduced performance of certain push strategies,

<sup>69</sup> Event hashtag predictor (retweet model) only approached conventional standards of statistical significance.

specifically those related to organisational information (ranked 7th) and non-social media content links (ranked 8th), as well as improved performances of certain pull strategy approaches, such as work recruitment (ranked 2nd) and interactive strategies (ranked 5th). The most significant observation is that service use is the most frequently retweeted strategy, which indicates that not only are support services being endorsed, but they are also being disseminated by digital audiences. In contrast, donations remain a low-ranking strategy, despite performing slightly better than networking in this model.

#### *6.4.2.5. Reply Model*

For content factors, only tweets related to celebrations and other focus areas showed no significant associations. In contrast, tweets containing interventions, a higher word count, and mentions were found to have a positive relationship with the number of replies, with no major discrepancies in effect size. In contrast, the presence of hashtags, COVID-19 mentions, and hyperlinks (effect sizes in this order). in tweets showed negative associations with the number of replies. These results suggest that tweets that are more interactive (i.e., using mentions to encourage dialogue from the outset) tend to receive more replies, whereas tweets that rely on external sources of information (i.e., hyperlinks), or use hashtags to reach a wider audience, may not always elicit dialogical engagement and interaction from their audiences. In relation to the tweet frequency reference category (high tweet frequency), low tweet frequency accounts were less likely to be replied to, whereas mid-tweet frequency accounts were more.

In relation to the account size reference category (x-large accounts), all remaining sizes were more associated with replies. The extent varied, with small accounts most associated replies, followed by large and finally medium. This finding suggests that smaller and more active organisations are more conducive to fostering an environment in which their digital audiences are inclined to participate in interactive engagements that involve a two-way exchange of information. No emoji predictors had significant associations with replies in the model. Time/ place hashtags had a positive association with replies. Interestingly, organisational hashtags were

negatively associated with replies, whereas branding hashtags were positively associated. From controls, day of the week had no significant associations with replies. In relation to the reference category (evening) tweets published in the morning and daytime were less likely to be replied to.

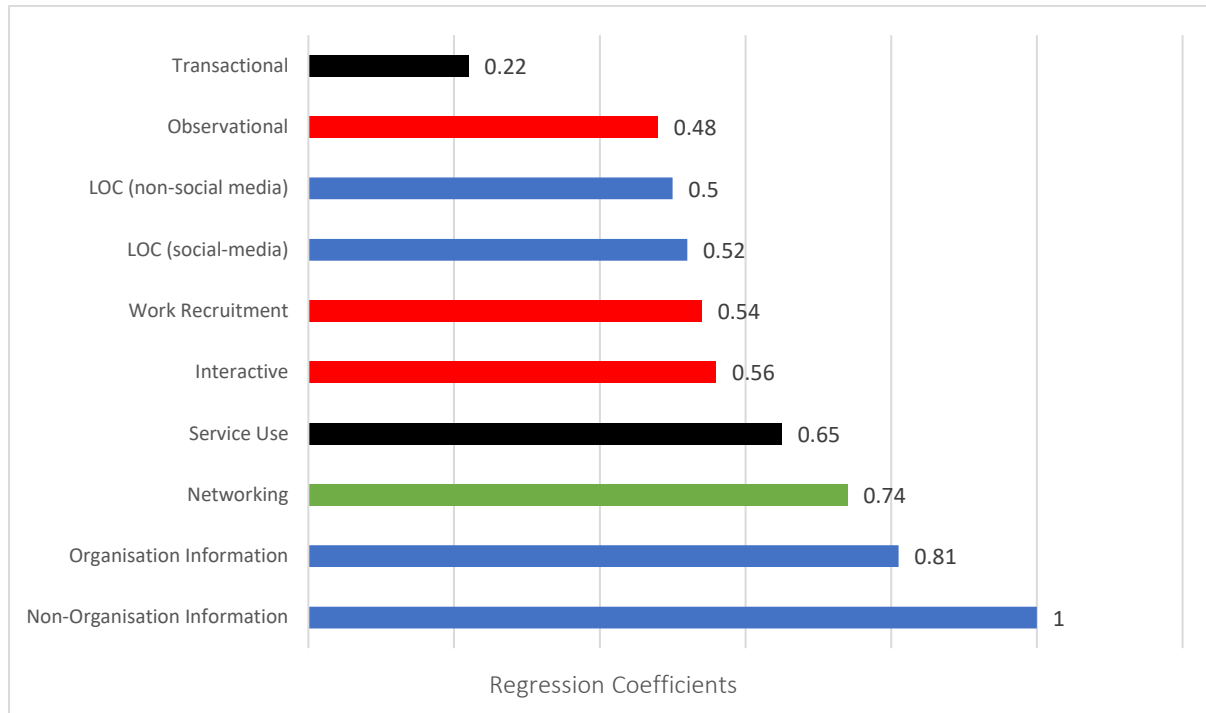


Figure 42: Audience engagement coefficients for 'replies' sub-strategy model

As mentioned earlier, in the broad strategy model both pull, and transactional strategies are less associated with reply count than push. Conversely, networking strategy tweets were most associated. The sub-strategy reply model (Figure 42) echoes this with significant departures from common themes seen in the like and retweets models. Although all eight remaining coefficients remain less associated with replies than the reference category (non-organisation information, rank 1<sup>st</sup>), many other sub-strategies rank differently. Perhaps most significantly is the increased likelihood networking strategy (rank: 3<sup>rd</sup>) has of being replied to. Despite, a much higher ranking this finding was expected, because the purpose of a networking communication strategy is to create an 'interactive' channel for organisations to discuss issues with followers in a constant 'back and forward' (Mergel, 2010). In fact, it is surprising that two push strategies (organisation information, rank 2<sup>nd</sup> and non-organisation information, rank 1<sup>st</sup>) outrank networking strategy in replies. This was

not previously detected in the broad models only relating to Mergel's (2013) framework. This finding indicates that although networking strategy tweets are designed to prompt online discourse with followers, particular push strategy tweets can still lead to more. Nonetheless, the significant difference between networking coefficients for replies and other engagement metrics, suggest that in third sector applications, it is an effective strategy for enhancing discussion with key audiences.

## 6.5. Discussion

Pro minority rights NGOs were found to employ a plethora of digital communication strategies to promote community cohesion. Push strategies, an informational, allowing organisations to broadcast information at any time without having to leverage the media. This is particularly important for minority rights organisations whose content may not be deemed “newsworthy” (Greenberg and MacAulay, 2009; Goldsmith, 2015). On the other hand, pull strategies promote engagement and participation, with organisations using followers and stakeholders as resources or information sensors while allowing for greater information depth (Mergel, 2012). However, Mergel (2012) contends that organisations do not typically employ pull strategies to stimulate discussions on social media platforms. Nevertheless, networking approaches move beyond initial two-way pull dialogical approaches towards an even more interactive framework that is two-way cross-boundary. Such interactions enhance social capital and enable a round-the-clock interactive environment in which organisations can both strengthen existing relationships with constituents and other stakeholders, while also gaining new ones (Mergel, 2013). Lovejoy and Saxton (2012) highlight the potential of social media in facilitating community building and the establishment of 'issue networks'. Public audiences' network with organisations, resulting in an interactive co-design of services. Finally, the 'transactional' strategy involves organisations utilising online communication to transact physical goods or services with external stakeholders or audiences (Mergel, 2012).

While high-level strategy overviews provide a broad perspective, they also encompass various subfunctions, enabling the exploration of more specific sub-

strategies. For example, push strategy consisted of organisation information, link to other content (non-social media), link to other content (social media), and non-organisation information. Pull strategy consisted of interactive, work recruitment, and observational. Transactional strategy consisted of appeals for donations and the provision of support services. Networking, however, could not be further sub-categorised into sub-strategies. This is consistent with related work as networking communications have been identified as decentralised (Meijer and Thaens, 2013). This meaning that the services and tasks completed in networking communications relate to a much wider variety of objectives that are often unique each time. This reflects the co-design of services promoted in the approach, in which organisation constituents have an equal say in the interactions (Mergel, 2013). As mentioned, there were significant discrepancies in strategy prevalence, content factors, and their success in audience engagement. These are considered throughout this discussion section to explore how NGOs use social media to promote community cohesion and reduce tensions.

In support of H1, evidence was found that NGOs directly contribute to creating and maintaining collective efficacy in digital communities. Collective efficacy is "the process of activating or converting social ties among communities to achieve collective goals, such as public order or control of crime" (Sampson, 2010: 802). In the context of digital society, the definition of community is transformed. Castells (1997) argues that digitalisation of social life has resulted in communities no longer being organised based on spatial or temporal dimensions. Instead, many communities now exist in digital spaces and differ from their offline counterparts in size and composition (Dutton, 1996). Regardless, scholars agree that, like in offline contexts, digital communities also need to restrict online transgressions and establish norms of behaviour (Postmes et al., 1998).

In support of past work (Perry et al., 2016) NGOs were found to use informal mechanisms (see Sampson, 2001), mainly promoting cohesion and trust through positive campaigns that support inclusiveness. In addition, informal social controls, such as actively challenging antagonistic speech with public interventions were also observed. Sampson et al. (1997) propose that informal social controls and promoting cohesion and trust are the primary mechanisms for achieving collective efficacy.

While existing research on digital collective efficacy has primarily focused on the former, typically in response to specific events (Costello et al., 2017; Ozalp et al., 2020), this study examined the day-to-day actions taken by NGOs to mitigate tensions and promote inclusivity through general campaigns that expose audiences to positive content as well as social threats.

The findings show that NGOs employed cohesion and trust strategies more frequently than interventions, as hypothesised<sup>70</sup>. Information depth and information reach (Van De Velde et al., 2015) were two vital forms of disseminating positive information used by these organisations for promoting positive content that is inclusive. From NGOs, this was achieved in both pull and push strategies, using information reach, with hashtags to enhance searchability and push strategies such as non-organisation information. The posts frequently featured positive and informative content, which not only debunked inaccurate information about communities but also provided an in-depth portrayal of community norms, thereby facilitating a better understanding of them. In this approach, the frequent incorporation of knowledge hashtags was a prevalent tool, defined as "any hashtag that seeks to help inform audiences on key concepts and terms" (Saxton et al., 2015: 161). Knowledge hashtags are particularly effective in increasing awareness of specific communities by educating followers on relevant terminologies and concepts (i.e., #hatecrime #lgbt and #womensrights). Additionally, these tools are used to raise awareness about potential dangers and threats that may affect marginalised communities. For instance, they can shed light on discriminatory policy changes or unjust societal practices that may disproportionately impact minority groups (i.e., #dabill and #womensrights). Such tools can help spark meaningful discussions and inspire positive actions by bringing attention to these issues. The use of these hashtags appeared to be effective in generating public endorsement, as indicated by the positive associations observed with likes. These hashtags were instrumental in propagating and amplifying such concepts (being associated with retweets), thus bolstering information reach across social media platforms. This facilitated the

---

<sup>70</sup> This is likely because interventions only come in response to specific incidents so are sporadic in nature.



dissemination of information, which contributed to heightening awareness and promoting mutual trust.

While effective, these were mostly confined to on-platform posts that allowed for wider information reach. Seeing as many community tensions can be deeply complex, textual restitutions on twitter (Han et al., 2019) can be limiting in providing detailed discussion on particular topics. To address this, NGOs have pursued strategies using hyperlinks to provide greater depth of information. These are predominantly employed through observational pull strategies, enabling organisations to directly disseminate information at greater depths by directing audiences to their other websites. Another approach involves utilising link to content (non-social media) push strategies, that include links to information published externally to the organisation. These tweets are typically published by news organisations or other NGOs. Examples are provided below:

*"How many BAME teachers are there in Wales?"  
#BlackHistoryMonthUK <https://t.co/6JsS6M6iJH>" --- Race Equality Frst*

The image shows a screenshot of a news article from BBC Wales. The main headline is "Black History Month: How can schools teach black history?". The article is dated 19 October 2020 and is written by Bethan Lewis. It features three main sections:

- Role models:** This section discusses Martha Holman, a geography teacher in Zimbabwe who was targeted by the late President Robert Mugabe's government. It includes a photo of her and quotes her saying, "I don't understand how that subject will impact on young learners when they don't have black teachers."
- 'Lack of understanding':** This section features a quote from Riah Andrews, an English teacher in Denbighshire, who says, "It's not necessarily that I've experienced any racism or prejudice," she said. "It's more a case of lack of understanding and awareness."
- How many BAME teachers are there in Wales?:** This section cites statistics from the Education Workforce Council, stating that 1.3% of teachers were from a black, Asian or minority ethnic background, and that only 54 teachers in Wales were identified as black/African/Caribbean/black British.

Figure 43: News article provided in Twitter link (4)

*Our manifesto, #BringUsOurRights, was led by disabled people, for disabled people. It highlights key areas including rights, employment and health. 68% of respondents to our survey didn't feel like their rights were being adequately enforced. Read more:*  
<https://t.co/CCgoiM6VJi> <https://t.co/A85b3LZaN2> --- **Disability Wales**

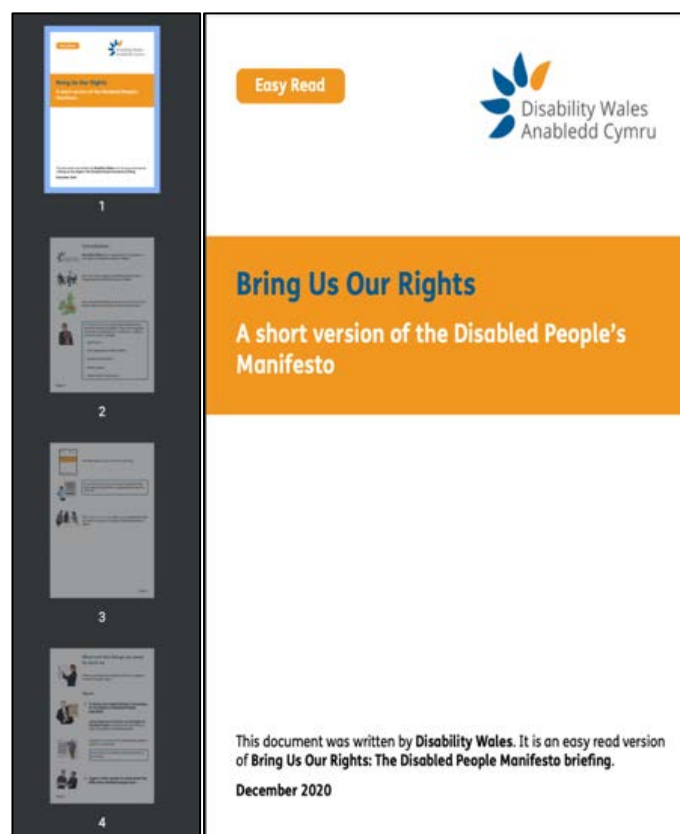


Figure 44: Organisation manifesto provided by Twitter link

These illustrate the value of including hyperlinks in social media posts to enhance the depth of information available to audiences beyond the limitations of character constraints. NGOs, for instance, can use hyperlinks to direct readers to additional resources that provide greater insight into other communities, fostering mutual trust and cohesion by improving understanding. Figure 44 shows this with the use of hyperlinks to direct audiences to the disability manifesto that outline critical experiences and key issues for affected groups. Link to other content approaches can direct audiences to very similar content to that published by the organisations, or in the example, full news articles that provide a more comprehensive understanding of important topics than just providing headlines. Hyperlinks facilitate longer format

discussions beyond the limitations of social media headlines. This approach provides audiences with a more detailed understanding of the issues at hand, which can foster meaningful conversations and mutual trust among diverse communities. The value of greater depth of information of other communities and groups lies in its ability to allow for greater exposure and understanding, bolstering cohesion and trust. Positive contact theory, as posited by Allport (1954), suggests that increasing exposure and knowledge of other groups can improve intergroup relations, resulting in collective efficacy (Sampson et al., 1997). In this way, hyperlinks can promote greater understanding and trust among diverse communities.

Informal social controls, or “intervention posts”, were highly (positively) associated with all three engagement models, indicating strong public endorsement for when organisations directly intervened and challenged discriminatory attitudes. This demonstrates that although these organisations' day-to-day operations are based on building cohesion and trust, the strongest endorsement comes when they intervene and directly challenge discriminatory attitudes. Perhaps the widespread endorsement of such interventions is bolstered by the general day-to-day build-up of community understanding and trust. As Simmons et al. (2005) put forward, they are not mutually exclusive, with the formation of tolerant spaces being critical in fostering an environment whereby communities are more likely to be involved or support public interventions. This is particularly true in interventions when organisations encourage the communities to mobilise collective responses such as public protests (Delli Carpini et al., 2004):

*"The protest is now on eventbrite for those who don't have facebook! <https://t.co/KxLFZJjlgm>" - **Trans Aid Cymru***



Figure 45: Protest event page provided by Twitter link

This demonstrates social media's potential to facilitate collective counter-speech by leveraging daily communities. These counteract injustices, including hate incidents, crimes, discriminatory policies, and related events in a concerted and cohesive manner. Specific protests and collective movements were enhanced by organisations using event and time and place hashtags that were both significantly endorsed across all three models, showing strong levels of support and reach for organised collective interventions. A major example of this was the global hashtag "#blackouttuesday" that has already been linked to informal social controls and collective efficacy by related work Gruver (2019). In a more local context, the "#LovePenallyHateRacism" hashtag was also used by many organisations, showing solidarity for refugees in the Penally asylum accommodation<sup>71</sup>.

To summarise, day-to-day exercises promoting positive content can help build a digital community through positive exposure that reflects Allport's positive contact theory (Allport, 1954) in which exposure to different communities bolsters mutual understanding and cohesion between constituents. However, it's essential to exercise caution and not overstate the findings of this study. The data may not unequivocally support the application of contact theory, as the research does not directly measure whether the posts increased interactions between in-groups and

---

<sup>71</sup> The extent to which collective efficacy manifested in this particular event was compressively explored in Chapter Five.

out-groups, nor does it assess whether this contact ultimately reduced community tensions. This aspect could be a focus of future research.

Provided these positive contacts are made, they could help build netiquette (Shea, 1994), fostering a sense of mutual understanding among digital constituents and establishing the norm that divisive behaviours are unacceptable, with the communities being more likely to intervene and signal when behaviour is unacceptable. In this vein, while not actively used for collective efficacy, networking strategies are also integral because collective efficacy requires a close-knit and supportive community. To achieve this, NGOs necessitate robust social capital (Dupont, 2004) with communities and external partners. H2 hypothesised that third sector minority rights organisations would utilise dialogical strategies on social media as an approach to try to enhance their social capital and networking within their communities and with external stakeholders. Indeed, posts containing dialogical hashtags were the most impactful in receiving replies to posts by a significant margin. Dialogic hashtags refer to “‘chat’ hashtags, prompting the audience to interact in discussion”<sup>72</sup> (Saxton et al., 2015: 161).

Other findings indicated a positive association between mentions - a prominent tool for dialogic content - and reply counts. This result is congruent with prior research highlighting mentions' effectiveness in facilitating online discourse and cultivating social capital (Burton and Soboleva, 2011). Consequently, mentions emerged as the most frequently employed networking strategy. This suggests that these organisations often employ networking strategies to establish relationships and partner with external stakeholders. Regression analysis supported this, showing networking strategy tweets were the third most predictive of receiving replies. Mergel (2013) posited that instant messenger functions on social media platforms are an extremely effective form of networking strategy. Nevertheless, due to understandable privacy restrictions, researchers are limited in accessing instant messenger data. This means that the prevalence data on networking

---

<sup>72</sup> Prevalent examples of this include examples such as ‘#askmecardiff’, ‘#haveyoursay’ and ‘#askme’.

communications does not include any private messages, leading to a potentially significant 'dark number' of unmeasured two-way interactive communications.

Contrary to the hypothesis, the two most associated strategies with replies were based in the broadcasting paradigm (Greenberg and MacAulay, 2009), namely, organisational, and non-organisational informational approaches, which made little to no use of dialogical approaches. These were primarily related to the community social capital aspect of H3, involving the development of relations with digital communities. Although these organisations used networking to engage communities in real-time communication on social media, they also employed diverse strategies for building an online presence, such as transparently showcasing their daily operations, similar to the image-building processes seen in the public sector (Mawby, 2010). These findings challenge conventional assumptions that one-way communications are typically tied to weak-tie audience relationships (Granovetter, 1973). This underscores the potential of social media harnessing transparency through broadcasting procedures, to establish closer, more trusted relationships with communities. This was reinforced by the positive association between branding hashtags and replies, indicating their effectiveness in generating online communication with communities. Branding hashtags are specific to an organisation, serving as unique identifiers for the organisation. These may include hashtags referencing the organisation's name and highlighting specific program names or support services offered by the organisation (Saxton et al. (2015: 161). Prominent examples of this include '#WomenConnectFirstStaff', 'BAMEHelpline' and '#PrideCymru'. Image-building strategies helps organisations to highlight their achievements and provide clarity about their daily operations, thereby augmenting their social capital, alongside building their searchability for new audiences (Kappeler and Gaines, 2020).

Organisations also utilise pull strategies, such as interactive approaches, to establish stronger relationships with communities. Interactive approaches help create a sense of community and facilitate events that serve purposes beyond information gathering, such as hosting quizzes or yoga sessions (Granovetter, 1973). These approaches utilised activity emojis to indicate specific events or actions, with the predominant use of symbols directly representing the activity, and occasionally

incorporating weather-related icons. The strategic implementation of these emojis demonstrated effectiveness in promoting organisational activities, eliciting online interest, as reflected by the positive associations with both likes and retweets. As such, activity emojis were primarily employed in the context of interactive pull approaches as illustrated in the results section.

Although activity emojis were successful in promoting community-based activities, surprisingly, no specific type of emoji was observed to facilitate two-way communication on social media platforms. Furthermore, no significant positive associations were found between the use of such emojis and replies. This finding contradicts previous research indicating that emojis effectively encourage less formal digital communication (Kwon and Sung, 2011). In addition to using activity emojis to stimulate community engagement in organisational community-building exercises, event-specific hashtags and time and place hashtags were used again to expand the audience reach and encourage participation. These hashtags served as essential tools in disseminating critical information about the events and helped to attract a broader range of participants. By leveraging these hashtags, organisations could effectively connect with wider audiences and facilitate greater engagement and participation in their community-building endeavours.

Smaller organisations were found to commonly receive the most public engagement. This suggests that scale is not necessarily important in building close-knit relationships with communities, but rather could be harmful in doing so. A potential reason for this could be communications between communities and small organisations using less formal language, which is more relatable to communities than the more formal language perhaps used by larger organisations. This conceptualisation supports related literature, such as Brainard and McNutt (2010), which suggests that using a professional tone in digital environments can make organisations seem less relatable and may disenfranchise audiences. However, it is important to note that this is merely based on visual interpretation by the researcher. Future studies could use a machine classifier to measure empirical differences in language formality between organisations based on factors such as size. In summary, the study aligns with previous research indicating that NGOs leverage social media to promote community and organisational involvement, thereby

fostering community cohesion (Vance et al., 2009). The findings further highlight the significance of social capital enhancement as a transformative outcome of social media use for such organisations. Beyond this, the study also reveals additional transformative features that social media can provide for these organisations.

H3 sought to explore how NGOs utilise social media to foster a transactional environment with their communities and investigate the prevalence and engagement with different transactional strategies. In support of H4, NGOs employ social media to foster a transactional environment, not only in soliciting donations, but also leveraging it as a platform to offer support services. This is consistent with previous research on transactional strategies in the public sector, where online customer service experiences are increasingly in demand (Poister and Thomas, 2007). Donation appeals manifested in many ways that align with previous research on non-profit transactional incentives, such as prize giveaways, trivia contests, and product sales (Castillo et al., 2014). Donations ranked very low in engagement across all strategies measured. However, it should be noted that the study only measured engagement on platform metrics, such as likes, replies, and retweets, without knowledge of click or donation rates, which would give a clearer picture of the effectiveness of donation strategies.

Interestingly, the study revealed that while soliciting donations was a common transactional strategy employed by NGOs, it was not the most prevalent. Rather, the promotion and provision of support services was more frequent and highly engaged with across all metrics. This suggests that social media can be an effective tool for providing support services, not only for hate victims but also for a wide range of issues. Scholars have traditionally emphasised the significance of offering support services to marginalised groups. Nonetheless, they have also voiced concerns about the limited accessibility of such services. This indicates that only a select few, often those in more stable economic or social positions, are aware of these services. Interestingly, even among these privileged individuals, many remain uninformed about such resources (Hardy and Chakraborti, 2020:144). The findings suggest a substantial level of engagement with support service approaches, pointing to the potential of social media in rectifying this issue by increasing the visibility and accessibility of these services to a broader audience. It's important to acknowledge



that not everyone has access to social media, so some individuals may still be excluded. Nevertheless, enhancing awareness through these platforms could prove beneficial in reaching a larger number of people. This finding is significant, as it has not been previously investigated in strategic communication frameworks and offers important evidence for future social media communication strategies. Phone icon emojis were frequently used tools for encouraging people to use support services. Furthermore, the study's findings support H3 and suggest that NGOs use social media to foster a transactional environment with their communities through soliciting donations and promoting and providing support services. Social media can be an effective tool for providing support services and should be considered as an important communication strategy for NGOs moving forward. This perhaps helps contribute to community building with online constituents that is pivotal for delivering other key aims such as upholding cohesion and collective efficacy.

Finally, in support of H4 NGOs used social media to gather information from the public and use them as 'sensors' to inform their initiatives. Organisations employed several approaches to gather information from the public. A common strategy was using interactive pull strategies to engage citizens and feed information back to organisations (Mergel, 2012). By promoting engagement and participation, organisations could gather valuable insights from the community, which could then be used to inform their operations. For instance, many organisations invited followers to participate in activities, focus groups, town hall meetings, or complete surveys designed by the organisation. These approaches not only resulted in enhancements of social capital with external groups or communities but also allowed organisations to better understand community members' experiences and opinions (Meijer and Thaens, 2013). In addition, the use of hyperlinks to digital surveys was also observed in some cases. However, many organisations used images to provide event details, including locations and start times, and did not require hyperlinks. Networking strategies also allowed for a two-way cross-boundary (Mergel, 2013) real time communication between communities and organisations whereby quick exchanges of critical information were possible, allowing for organisations to receive critical information in real-time that could not be achieved at the same speed with interactive pull activities such as scheduled focus groups. Overall, NGOs utilised various dialogical strategies, such as interactive pull and networking approaches to

gather information from the public and use them as sensors to inform their initiatives. By actively engaging with citizens and promoting participation, organisations could gain insights that could guide their day-to-day operations.

## **6.6. Summary**

This chapter has explored how minority rights NGOs use communication strategies on social media to address social injustices and promote community cohesion in both online and offline contexts. While a wide body of literature extensively examines the work of organisations in offline contexts, there has been comparatively less exploration of their online activities. Early research findings indicate that social media can be a fruitful avenue for addressing community tensions and subsequent associated issues such as hate speech and discriminatory attitudes (Chakraborti, 2018:400). This study expands our understanding of how organisations utilise social media and digital content aiming to tackle these challenges by exploring the various strategies they employ to harness the potential of social media. The study found that organisations employ many strategies to leverage social media, each serving a distinct purpose while ultimately interlinking with one another. One way is through knowledge and awareness campaigns that aim to educate and inform the public about issues. By sharing information and resources, these organisations can raise awareness and promote understanding of discriminatory attitudes, whether below or above a criminal threshold on individuals and communities. Additionally, organisations can participate in and mobilise interventions where unacceptable behaviour or actions are challenged. These interventions can range from responding to intolerant content and misinformation online to organising community events and protests that aim to promote tolerance and inclusivity. Both knowledge and awareness campaigns and interventions feed into collective efficacy (Sampson et al., 1997), empowering individuals and communities to act against discriminatory and, or misinformed attitudes.

Another way minority rights NGOs use social media is by building relationships with their audiences and other relevant stakeholders. By doing so, they may be able to foster a sense of community among their followers and create a safe space where

people can share their experiences and concerns. This, in turn, helps to strengthen cohesion and collective efficacy, which are important for addressing social injustices. The study also found that these organisations use social media as an opportunity to foster a transactional environment with their communities. They do this not only through soliciting donations but also through promoting and providing support services. Social media can serve as an effective tool for providing support services, and it should be regarded as an essential communication strategy for third sector organizations in the future. It enables previously hard-to-reach constituents who are unaware of services (Hardy and Chakraborti, 2020) to access them more easily. Finally, minority rights NGOs use social media to gather information from the public and use them as 'sensors' to inform their initiatives. They employ various dialogical strategies, such as interactive pull and networking approaches, to gather insights to guide their day-to-day operations. Organisations can gain insights that inform their activities by actively engaging with citizens and promoting participation. Overall, the findings suggest that minority rights NGOs use social media in many ways to attempt to address social injustices and promote community cohesion. While each strategy serves a distinct purpose, they interlink with one another to create a comprehensive approach to combating community tensions. While this demonstrates a multifaceted approach, this thesis makes no assumptions about its ultimate effectiveness, which could be measured in follow up research.

# Chapter 7

## Discussion

---

### 7.1. Introduction

This chapter explores the key findings from the three empirical investigations conducted in the previous chapters and brings them together into overarching themes. The sections covered in this chapter include External Mechanisms of Cohesion Delivery, Informal Mechanisms of Cohesion Delivery, The Third sector: A Hybrid Approach to Collective Efficacy and Cohesion Delivery in Digital Society. Each discusses the pursued ways of trying to fostering unity and reducing tensions, mostly grounded in the conceptual framework of collective efficacy and digital society. Drawing on the synthesis of all three empirical investigations, this chapter provides valuable insights into the various mechanisms used to try to facilitate cohesion delivery in contemporary Wales. This discussion chapter synthesises the research findings from the three empirical chapters to address the central research question of this thesis - "How is community cohesion attempted to be maintained and promoted in contemporary Wales?". The multifaceted nature of this question is highlighted, with cohesion being addressed by a diverse range of social actors across various sectors, including within communities themselves. The discussion explores the role of emerging technologies in digital society and persistent threats to cohesion delivery in modern Wales.

### 7.2. External Mechanisms of Cohesion Delivery

As discussed earlier in Chapter Two, an important aspect of collective efficacy is the presence of externally induced actions, known as formal controls (Sampson et al., 1997, p. 918). These were explored multiple times to gain insights into their impact on collective efficacy. As scholars note, traditionally in many formal approaches to policy issues, externally induced actions were predominately centralised to public

sector nodes (Kickert et al., 1997). However, the turn of the millennium has shifted towards a more pluralised approach involving a wider range of social actors (Crawford, 2002). This shift acknowledges the importance of involving diverse actors and fostering partnerships to address social issues and promote collective efficacy (Bayley, 2001).

### ***7.2.1. Nodal Governance Multi-Agency Partnership***

The findings of this thesis provide compelling evidence of a nodal governance approach being implemented for cohesion delivery in Wales. Nodal governance entails collaborating and coordinating various stakeholders from different sectors, including the voluntary, private, and community sectors, to tackle social issues (Burris, 2004). This approach encourages the diffusion of responsibilities among multi-agency networks of social actors who previously lacked regulatory mandates (Levi & Williams, 2013).

Based on the typology proposed by Blake et al. (2008), this thesis explored the extent to which a nodal governance system is employed, considering nodes such as central government, the police, regional teams, civil society, evidence practitioners, and communities themselves. While the literature on how these partners uphold collective efficacy through a cohesion lens is limited, there has been more discussion on how they interact with collective efficacy at a general level. It is important to note that all these partners, except community groups, operate through externally induced actions (Sampson et al., 1997).

Chapter Four of the thesis presented comprehensive evidence of the implementation of nodal governance for cohesion delivery in Wales. Initial indications laid out in the cohesion agenda set by the Welsh Government demonstrate clear intentions to incorporate many actors and perspectives, fostering a multi-agency approach that capitalises on their unique contributions (Burris, 2004). Analysis conducted through Multidimensional Scaling (MDS) and Principal Component Analysis (PCA) procedures offered the most salient evidence confirming a wide involvement of diverse stakeholder groups in cohesion delivery, further supporting the evidence of

nodal governance. The nodal governance systems observed in the Welsh CDN build upon the benefits derived from the multi-lateralisation reforms of new public management. Multi-lateralisation allows for the specialisation of partners in different areas and the integration of their unique knowledge and capacities (Johnston & Shearing, 2003). As conceptualised by Dupont (2004), cultural capital emphasises the specialised knowledge of different network nodes. Findings in Chapter Four highlight the diverse forms of cultural capital brought by different stakeholders in the CDN, which enhance collective efficacy in reducing tensions.

### ***7.2.2. Cultural Capital and Specialist Roles of Network Stakeholders***

Regional cohesion teams, comprising both cohesion officers and coordinators, were found in the open output to work at the 'ground level', to address community cohesion at a local basis. In this vein, cohesion officers are designed to serve as a valuable link between communities and the authorities, providing insights and understanding of specific tensions and challenges that are not always visible at the higher level. They are responsible for engaging with local communities and to try and build trust, with the aim to develop tailored interventions that consider the community's unique cultural and social context. Cohesion coordinators, on the other hand, operate at a higher level, developing and implementing wider strategies for promoting cohesion and community resilience. They are responsible for creating a cohesive and inclusive society by developing policies, programs, and initiatives informed by the insights gained by cohesion officers on the ground. They also work in digital spaces to better understand their communities, particularly those with a strong online presence or use social media extensively. For instance, cohesion officers have been instrumental in resolving community tensions and incidents related to race, religion, and ethnicity by providing a nuanced understanding of the community's specific cultural and social context. In addition, open output in Chapter Four revealed cohesion coordinators have developed policies and initiatives that aim to promote social inclusion and cohesion, such as community-based initiatives for youth, interfaith dialogues, and cultural exchange programs.

The police are an essential component of the multi-agency network that aims to promote community safety and well-being. They possess specialised knowledge in law enforcement, community safety, and conflict resolution, contributing to the network's external influence on collective efficacy. By enforcing the law, they could create a normative climate that promotes social order and cohesion within the community. Community-oriented policing approaches are a key strategy the police uses to try and promote social harmony. This approach involves working closely with community groups and other agencies to identify and address community safety issues. Through active engagement with the community, the police can build trust and promote a sense of shared responsibility for community safety. The police's cultural capital, including their ability to understand and navigate the community's diverse cultural and social contexts, can play a vital role in promoting cohesion. For instance, they can work with community leaders to promote intercultural understanding and respect between groups. They can also use their powers to enforce the law and create a normalised feeling that many behaviours driven by these biases are unacceptable. However, it should be noted that the police's role in promoting community safety and cohesion is not without controversy. Some scholars argue that the police's use of force and other coercive tactics can undermine trust and exacerbate tensions between the police and the community. Others argue that the police's cultural capital can be limited by the perception of institutional biases. This will be further explored later in this Chapter.

The Welsh government is crucial in the approach taken aimed at promoting regional cohesion by leveraging its cultural capital. It sets the cohesion agenda, provides guidance, resources, and strategic direction, and fosters stakeholder collaboration. This aligns with the concept of political capital, as defined by Dupont (2004), which empowers nodes to influence policies, allocate public expenditure, and strategise multi-agency responses. The government's endorsement of a nodal governance system, as outlined in the framework for action (2017), serves as testimony to its commitment to improving the collective response to community tensions in Wales. The framework outlines four central objectives: prevention, reporting, support, and enforcement, and multiple sub-delivery areas. For example, under the prevention objective, the government focuses on improving the understanding of underlying causes of tensions,. Under the reporting objective, the government focuses on

improving the reporting mechanisms, increasing the number of reported incidents, and working with the police to improve data collection and analysis. Previous research emphasised the significance of third-party organisations in improving hate crime reporting (Schweppe et al., 2020). This study extends that perspective by recognising the potential of such groups, specifically NGOs, to not only collect hate crimes but also incidents falling below the criminal threshold, such as microaggressions, to allow for a more comprehensive picture of the landscape when addressing community tensions. This potential can be further optimised through social media monitoring processes, which will be discussed later in this chapter. Under the support objective, the government supports the victims, addresses their needs, and ensures that they receive appropriate services. Finally, under the enforcement objective, the government improves the response to hate crimes, increases the number of successful prosecutions, and raises public awareness about the consequences of hate crimes and wider tensions.

The Welsh government's active involvement in the network reinforces the nodal governance approach and emphasises its commitment to enhancing the multi-agency cohesion delivery. By leveraging their political capital, providing strategic direction, and fostering stakeholder collaboration, the Welsh government plays a pivotal role in shaping policies, allocating resources, and strengthening the CDN. Their active involvement further demonstrates their commitment to addressing community tensions through a multi-agency approach. The government's efforts strengthen collective efficacy in addressing tensions, thus promoting a more cohesive and inclusive society.

NGOs are uniquely positioned due to their proximity to grassroots organisations and individuals, allowing them to possess a cultural capital that gives them a deep understanding of communities' specific needs and challenges. This understanding is based on their strong social capital with communities (Dupont, 2004). Building on past work, (Perry et al., 2016) Chapter Six illustrates, NGOs can be critical in leading knowledge and awareness campaigns to address and combat misconceptions. Their practical knowledge and flexible approaches enable them to address emerging issues and promote cohesive communities. We will explore the third sector's role in greater depth later in this chapter.



The research findings highlight that different stakeholders in the Welsh CDN bring their own unique forms of cultural capital to the partnership. This includes knowledge, skills, and experiences related to different cultural backgrounds, traditions, and practices. By working together, stakeholders can leverage their diverse cultural capital to promote community cohesion and enhance the overall collective efficacy of the partnership. This underscores the importance of recognising and valuing different stakeholders' contributions and cultural resources in creating a more inclusive and cohesive society. Realising the benefits associated with specialised knowledge, or cultural capital, within multi-agency networks, is contingent upon maximising cooperation patterns (Levi & Williams, 2013). In this context, Dupont's (2004) concept of social capital assumes a crucial role, emphasising the importance of effective communication and cooperation among network nodes (Halpern, 2005). By fostering robust cooperation, multi-agency networks can unlock and leverage different stakeholders' unique expertise and knowledge, thereby enhancing collective efforts in addressing critical issues. These were further explored in the MDS and PCA cluster analysis procedures. The analysis identified the presence of social capital hotspots and faults within the CDN, forming three distinct stakeholder clusters. These clusters included: (1) regional agents, consisting of the police and cohesion teams; (2) third sector agents, comprising NGOs and community groups; and (3) policy/evidence agents, involving the Welsh Government and academics. Importantly, no stakeholder group was entirely isolated from the network, as all stakeholders contributed to a meta-cluster, providing further support for a pluralised approach.

### ***7.2.3. Network Cooperation Patterns***

The existing literature has emphasised the importance of increased cooperation among all network stakeholders to achieve effective outcomes over the turn of the millennium (Kean & Hamilton, 2004). However, this thesis challenges this notion by revealing that optimising cooperation does not necessarily require increased communication among all nodes. Instead, the study findings indicate that specialists, such as academics, can make valuable contributions to the network by providing empirical-driven, evidence-based information that offers valuable insights into the

dynamics of social tensions and the effectiveness of cohesion strategies. Their specialised knowledge enables the network to make informed decisions and implement targeted interventions that foster collective efficacy at multiple levels. Moreover, academics possess a strong informational cultural capital (Dupont, 2004) that can aid in developing evidence-led approaches. However, the study participants, including academics, indicated in the open outputs that their regular involvement in the network was unnecessary, and their inputs could still hold importance when given periodically. They viewed the Welsh Government as a "pressure point" (Legg, 2021) for knowledge exchange as it emerged as a gateway for diffusing academic cultural capital into the network. This aligns with Posner's (2009) perspective that governments often involve academics in networks due to their vast knowledge and expertise. Studies on addressing community cohesion have found that academic involvement in partnerships can often be tokenistic and merely a box-ticking exercise (Chakraborti, 2016). However, the findings for the Welsh CDN indicate that even though their cooperation frequency may not be very high, their cultural capital and expertise are effectively diffused into the network through periodic but meaningful interactions (Dupont, 2004). Therefore, the study recommends that network managers should recognise and prioritise the importance of specialised knowledge and expertise and involve specialists like academics in the network, even if not regularly. By adopting this approach, other networks across the UK and globally can learn from the Welsh CDN and move towards a more evidence-based approach, where academic engagement is integrated in a more meaningful way, prioritising the quality of interactions over quantity.

Although uncommon, cooperation gaps were observed within intra-cluster nodes, particularly regarding horizontal information-sharing between cohesion teams and the police. Cohesion teams suggested in the open outputs that much of the information gathered by the police relating to incidents below a criminal threshold were not regularly shared. Such sentiments were reflected in the quantitative quality of cooperation rankings that showed much lower perceived quality from cohesion teams' perspectives towards the police than visa-versa. This indicates a collaboration gap resulting from non-reciprocal communication. This finding aligns with Cheminais' (2009) argument that poor information-sharing culture can weaken multi-agency partnerships, reducing the quality of cooperation. Despite

acknowledging the benefits of localised partnerships, regional stakeholders such as the police and cohesion teams still encountered challenges related to information-sharing. One potential reason identified in the open output across both stakeholder groups for this was a lack of clarity in the overall multi-agency agenda set by the Welsh Government. Perhaps by reevaluating the CDN framework, the Welsh Government, with its political capital (Dupont, 2004), could establish clearer indicators for stakeholders regarding what, how, and when to share information with other partners, thereby minimising duplication faults common in multi-agency networks (Cheminais, 2009) and enhancing cooperation and productivity. Addressing such issues is paramount to optimising all stakeholders' contributions to collective efficacy.

#### ***7.2.4. The Role of The Media in Cohesion Delivery***

This thesis acknowledges the central role of network nodes in reinforcing cohesion, but it also recognises the significance of other external formal bodies, particularly the media<sup>73</sup>, in the potential of achieving cohesion and collective efficacy. The media was predominately measured throughout this thesis based on news headlines, and therefore readers should keep in mind that other sources could exhibit notable differences, The findings presented in Chapter Five demonstrate a positive association between press headlines and overall information dissemination relating to the Penally asylum accommodation. These results align with the notion that the media continues to shape public discourse, particularly following significant events, by setting the agenda and transmitting images (Cohen, 2002).

In the specific case of the Penally asylum accommodation, the media were vital in influencing information flows, primarily through conventional approaches such as newspaper publications rather than newer media platforms like Twitter (i.e. news actors, such as the BBC were not associated with overall information dissemination when tweeting directly) These findings are encouraging, revealing the media's

---

<sup>73</sup> The media was predominately measured by exploring written news headlines and readers should note other media forms could exhibit notable differences.

potentially more neutral role in such events, while retaining influence over information flows. When the media report accurately and impartially, they have the potential to leverage their influence debunk misinformation and contribute to collective efficacy by utilising symbolic capital (Dupont, 2004) to challenge divisive attitudes, reinforce intergroup trust, and foster mutual understanding (Sampson et al., 1997).

Additionally, the findings discussed in Chapter Six highlight that knowledge and awareness campaigns conducted by NGOs often leverage other media outlets to try to support and legitimise their points, aiming to promote trust, understanding, and cohesion. This included links to articles published by institutions such as the BBC that supported or provided information about issues relating to minority rights campaigns. Thus, the media has retained significant influence, particularly in the dissemination of information around community tension issues. When employed appropriately, the media can play a crucial role in achieving collective efficacy. The findings demonstrate their potential to shape public discourse, that if used correctly could assist to debunk misinformation, challenge negative attitudes, and contribute to the establishment of intergroup trust and understanding. By harnessing the influence of the media through accurate and unbiased reporting, initiatives aimed at achieving collective efficacy have the potential effectively leverage symbolic capital to address community tension issues and promote cohesion. While findings in this thesis reveal their continued influence, it makes no assumptions in their ultimate effectiveness in achieving, or not achieving, cohesion.

### **7.3. Informal Mechanisms of Cohesion Delivery**

#### ***7.3.1. Introduction***

In contrast to externally induced actions, informal mechanisms of collective efficacy encompass processes that occur within the community, facilitated by informal constituents. These mechanisms can be broadly classified into two categories: "cohesion and trust" and "informal social control" (Sampson et al., 1997, p. 918). Research, as indicated by Sampson et al. (1997), has demonstrated that these

control mechanisms are not mutually exclusive, as further evidenced by the findings in this thesis.

Cohesion and trust, as informal mechanisms, are fostered when residents gain a deeper understanding of the various groups residing in their community. This understanding allows for stronger bonds of cohesion and trust that transcend previous feelings of threat and strain (Sampson et al., 1997). Establishing common norms, values, and shared interests among sub-communities contributes to forming these bonds (Ferlander & Timms, 1999; Forrest & Kearns, 2000; Lynch, 2001). The degree to which communities uphold cohesion depends on mutual trust, understanding, and tolerance within a society. Mutual trust can serve as a preventative measure to uphold collective efficacy.

In contrast to the proactive nature of cohesion and trust, informal control can be a reactive mechanism in pursuing collective efficacy (Warner, 2014). As defined in contemporary perspectives, informal social control encompasses a range of actions citizens undertake to indicate and address unacceptable behaviour (Groff, 2015, p. 90). Within the framework of collective efficacy, these actions entail citizen interventions aimed at advancing the common good within their community (Sampson, 2012). Informal social controls can be further classified as "direct" or "indirect" (Warner, 2007, p. 99). While informal social control is typically seen as a reactive approach, it is important to note that when individuals signal and intervene against unacceptable behaviour, it can cultivate mutual trust and understanding among constituents (Cantle, 2018). Therefore, informal social control addresses negative occurrences and is vital in promoting positive social dynamics, including developing mutual trust, and understanding within the community.

### ***7.3.2. Impact of Informal Mechanisms and Real-World Evidence***

Chapter Five provided real-world evidence of direct informal controls in practice in the context of the announcement of the Penally asylum accommodation trigger event. Consistent with previous work, constituents frequently engaged in counter-speech, actively challenging negative attitudes directed towards refugees (Ozalp et

al., 2020). This reflects the belief within the community that its members can effectively collaborate to address social issues. These instances were often observed in the highly replied posts that were also the most antagonistic in the dataset. These actions publicly signalled that such sentiments were unacceptable (Groff, 2015). The community's involvement demonstrated collective efficacy, as individuals recognised their power to effect positive change and combat discriminatory attitudes. By speaking out, they displayed a shared belief in their ability to create a more inclusive and accepting community. While there was evidence to show they mobilised to challenge divisive attitudes, the thesis did not measure its overall effectiveness, a compelling area of focus for follow up studies. Informal mechanisms of collective efficacy were observed through various actions, such as pro-refugee activism and the use of counter-speech against anti-refugee activists and far-right political agents. Community members leveraged their social influence, personal convictions, and connections with external bodies, such as the police, to challenge negative attitudes and promote empathy towards refugees. These actions exemplified the strength of community bonds and shared values, driven by a sense of responsibility and commitment to support causes aligned with their collective identity (Sampson et al., 1997). Engaging in these informal mechanisms of collective efficacy challenged negative attitudes and could possibly foster a sense of empowerment among community members. By actively countering and normalising the rejection of discriminatory attitudes, they contributed to the well-being and cohesion of the community (Cantle, 2018). Furthermore, these actions strengthened social connections, encouraged solidarity, and perhaps reinforced the shared responsibility to create a more inclusive society after a trigger event. It should be noted that the effectiveness of this in action was not directly measured. These findings contribute to the existing literature by demonstrating that trigger events not only mobilise discriminatory attitudes (Williams and Burnap, 2016) but also have the potential to act as catalysts for positive counter-movements, possibly leading to increased solidarity, acceptance, and collective efficacy within communities.

### ***7.3.3. Community Participation with External Stakeholders***

The second facet of informal social controls is indirect, which involves communities mobilising and interacting with formal authorities to address community tensions. Sampson (2001) argues that while informal mechanisms stem from community actions in informal capacities, the level of integration between communities and external forces is also crucial for achieving collective efficacy at informal levels. The consequences of poor integration of communities into multi-agency partnerships can be far-reaching. Inadequate integration can lead to limited awareness of policies and support services, undermining their effectiveness. It poses a significant threat to the ability of networks to detect, respond to, and prevent incidents, crimes, and underlying tensions, as research has shown (Chakraborti, 2018; Garland et al., 2022), potentially causing a lack of integration can result in duplication of efforts, miscommunication, and a lack of coordination among network members. Ultimately, this may harm the community, as incidents and underlying tensions may go undetected or unaddressed.

Furthermore, community members may become disempowered to mitigate tensions by actively challenging it or calling external partners into action (i.e., the police) (Sampson et al., 1997). Therefore, achieving collective efficacy and mitigating tensions depends on achieving a suitable and meaningful integration of communities into external multi-agency partnerships. This integration is achieved through meaningful community participation with external agencies and the establishment of multi-agency partnerships. To conceptualise the extent of engagement between informal partners, such as community members and groups, and external partners, Arnstein's (1969) "Ladder of Engagement" was considered. This provides a valuable framework for assessing the level of citizen participation, or denizenship, within multi-agency networks. It presents a spectrum of participation, ranging from forms that fall significantly below meaningful implementation to those demonstrating varying degrees of tokenism and citizen power. The ladder consists of eight distinct rungs that can be categorised into three broader classifications of citizen participation: non-participation, tokenism, and genuine citizen power.

Community participation was assessed in Chapter Four, which focused on exploring the relationships between stakeholders and community groups. Incorporating multiple actors within the cohesion agenda aims to facilitate a multi-agency approach (Burriss, 2004). These intentions align with Arnstein's concept of 'partnership', which suggests that community groups possess a degree of citizen power to influence network development and agenda propositions. However, analysis revealed that the desired levels of integration may not have been fully realised. The current state of community engagement within the CDN appears to exhibit a 'degree of tokenism' rather than a 'degree of citizen power' (Arnstein, 1969: 217), indicating a lack of meaningful citizenship (Shearing and Wood, 2003) within the CDN.

Currently, this mirrors the 'placation' stage on Arnstein's ladder, where citizens are afforded only limited opportunities for influence. This is evident in their peripheral placement in the MDS visualisation and their consistently low cooperation rankings. These findings raise concerns as meaningful citizen participation is essential for fostering collective efficacy and community cohesion. To overcome tokenism and enhance community cohesion, it is crucial to establish avenues for genuine engagement, such as town hall meetings where inputs result in tangible influence, allowing community groups to play a meaningful role in decision-making processes and contribute to policy and practice development (Blake et al., 2008). This emphasis on active citizen participation and engagement strategies is essential to strengthen community cohesion and achieve meaningful citizenship within the CDN. Tokenism, characterised by superficial involvement, fails to provide substantive citizen partnership and input. This superficial engagement hinders the development of genuine collective efficacy and collaborative efforts within the community.

Trust is pivotal in bridging gaps and achieving collective efficacy by fostering collaboration, open communication, and a shared sense of responsibility between external partners and communities (Cantle, 2008). This emphasis on trust could create essential spaces for dialogue, where community groups can actively contribute to network planning and decision-making processes, thereby strengthening community cohesion (Burriss, 2004). Results in Chapter Four revealed significant issues stemming from poor cooperation between the police and



communities, which poses challenges to upholding collective efficacy. Extensive literature exists on the role of the police in maintaining collective efficacy, with studies examining their actions as external agents in this regard (Hart & Colavito, 2011; Nix et al., 2015; Wiesburd et al., 2015). In the context of community cohesion, this entails the police enforcing hate crime laws and diffusing tensions, often relying on communities reporting issues to them.

The police's ability to externally influence collective efficacy emphasised the importance of legitimacy as a critical factor for success (LaFree, 1998). Establishing trust and legitimacy within the police force allows them to uphold symbolic capital and improve relationships with the community (Dupont, 2004). In this context, legitimacy is rooted in the notion that agencies like the police have an obligation to intervene beyond self-interest (Kochel, 2012). The literature over time consistently suggests that the tactics employed by law enforcement significantly impact community perceptions of their legitimacy, trustworthiness, and effectiveness, consequently influencing the fostering or hindering of collective efficacy (Scott, 2002; Renauer, 2007; Sargeant et al., 2013). Therefore, the findings from Chapter Four such as low cooperation scores, particularly in the quality metric, as well as open outputs both from community groups, as well as charities indicating that communities are unwilling to cooperate with the police in favour of the third sector, or not at all are concerning. They indicate low levels of trust between the police and communities, suggesting a lack of symbolic capital for the police. The low levels of trust between the police and minority communities could be explained by factors such as historical discrimination, perceived lack of transparency, and ineffective community policing practices (Chakraborti & Garland, 2003; Pass et al., 2020). Low levels of trust between the police and minority communities can hinder meaningful community engagement, impeding efforts to improve community cohesion. This disconnect can lead to reduced reporting, cooperation, and shared responsibility, potentially perpetuating divisions and preventing inclusive community-building.

The cohesion teams, comprising coordinators and officers, emerged as nodes that demonstrated higher levels of symbolic capital with both higher cooperation scores in Chapter Four and direct inputs given in the open output, indicating greater levels of perceived trust and credibility from communities. This provides supporting evidence

for current recommendations advocating the integration of third-party (non-police) organisations, particularly those with substantial symbolic capital, into the reporting process, aiming to contribute to improved accuracy (Schweppe et al., 2020). Although the current level of community involvement in the cohesion programme may not align with initial intentions, it is important to consider the relatively recent development of community engagement roles, specifically the cohesion teams. While these teams and community groups currently exhibit moderate levels of engagement, data on cooperation quality suggests that as their roles mature, they will enhance community integration into the CDN, serving as effective gateway pressure points (Legg, 2021).

Thus, they have the potential to serve as effective gateways for community groups to enter the community development network, allowing for more meaningful and empowered citizen participation in community justice, akin to a degree of citizen power (Arnstein, 1969). Despite this potential, many communities that could benefit from the input of cohesion teams are unaware of their existence. This lack of awareness highlights significant gaps within the network. The limited awareness of these roles can be attributed to the relatively short time that officers have been in their positions. Therefore, careful consideration should be given to the further development, recognition, and establishment of cohesion roles to bridge the gap between traditional agencies and communities. By doing so, the network can ensure the meaningful and non-tokenistic involvement of communities, which are the primary stakeholders in cohesion delivery.

Contrary to the findings presented in Chapter Four, the findings discussed in Chapter Five reveal a stronger level of community participation with external partners. This is exemplified by the endorsement received by political and police agents, who were negatively associated with anti-refugee attitudes when discussing the Penally asylum accommodation. These endorsements indicate a higher level of symbolic and social capital between these external agents and the communities, surpassing the levels observed in Chapter Four. This was evident in the endorsement metrics, such as likes and retweets in Chapter Five. Moreover, they serve as strong indicators of a high degree of collective efficacy, as the communities endorse their positive sentiments. This expands on recent evidence that suggests capable, trustworthy, and willing external actors can effectively counter negative attitudes by mobilising

general population support, thereby enhancing collective efficacy and reducing tensions (Ozalp et al., 2020).

This suggests that when organisations working to counter harmful narratives, such as antagonistic speech, become active on social media platforms, their messages propagate further and have a longer-lasting impact than antagonistic messages. This finding underscores the significance of the work carried out by organisations dedicated to safeguarding communities and promoting collective efficacy. One potential reason for the increased endorsement from communities towards public sector entities, such as the police, is their presence on social media platforms. This aligns with recent evidence suggesting that digital society has transformative potential (Powell et al., 2018) for institutions historically lacking symbolic capital (Dupont, 2004), enabling them to establish new and stronger relationships with communities through social media platforms. By leveraging these platforms, these entities can enhance their digital image and become more accessible, thereby improving their relationship with the public (Mawby, 2002). The implications of digital society on the delivery of cohesion will be further explored later in this chapter.

While these endorsements perhaps demonstrate that communities may have more trust in traditional public agencies, such as the police, than initially anticipated in chapter four, it is important to note that these interactions, particularly in digital contexts, tend to be predominantly one-way. There is limited evidence of public sector entities encouraging a meaningful two-way approach whereby communities are genuinely integrated into their operations. From this perspective, citizen partnership remains mostly tokenistic (Arnstein, 1969), failing to reach a point of denizenship where communities contribute meaningfully to operations (Shearing and Wood, 2003). Disparities in community engagement with external formal bodies have been observed, highlighting varying degrees of involvement. However, a consistent pattern of meaningful interactions was identified between community groups and the third sector throughout the study. Consequently, this pattern, along with other influential factors, sparked a heightened interest in the role of the third sector in fostering collective efficacy.

## **7.4. The Third Sector: A Hybrid Approach to Collective Efficacy**

### ***7.4.1. A Catalyst for Community Engagement***

This thesis investigated the third sector's contribution to collective efficacy in various contexts. The initial indication of the third sector's significant role emerged from examining its relationship with communities. As mentioned earlier, informal community mechanisms are crucial for collective efficacy, but the level of integration between informal actors and external organisations is paramount for its success (Sampson, 2001). The findings across the three chapters consistently demonstrated a high degree of social capital shared between NGOs and communities.

Chapter Four provided the first evidence of these strong social ties. They were closely associated, loading into a single correlated meta-cluster with an exceptionally high cooperation frequency. This level of cooperation surpassed other stakeholder groups, except for cohesion team sub-clusters, whose roles were formed together and naturally necessitated close collaboration. Subsequent analysis of the quality of cooperation further reinforced strong relations. While cohesion teams were previously discussed as potential gateways for community input into the network, this was based on speculation and the assumption that as their roles become more well-known among communities, cooperation and integration into the community development network would increase, enhancing the level of denizenship (Shearing and Wood, 2003). However, the exceptional social capital between the third sector and communities indicates that they currently serve as a crucial access point to the network (Legg, 2021). Respondents in Chapter Four noted that NGOs are more approachable, presenting a less formal and more amicable appearance, reflecting their capability and trustworthiness, thereby indicating high levels of symbolic capital (Dupont, 2004). In contrast, other stakeholders, such as the police, are actively avoided by certain community members.

Another noteworthy finding that helps explain this close relationship is the tendency of NGOs to specialise in addressing specific group characteristics, such as disability, LGBT+, race, and ethnicity. Communities perceive this specialisation as more beneficial in addressing their specific concerns and providing targeted support than

other stakeholders who adopt a more generalised approach. While encouraging, current data indicates a tendency to broadly group that share similarities. Stakeholders often combine characteristics such as gender reassignment and sexual orientation, potentially neglecting their specific needs as cautioned by Haynes and Scheppe (2017). Despite this, the exceptionally strong social capital shared between the third sector and communities represents the primary avenue through which the third sector contributes to collective efficacy.

Chapter Six delved deeper into how this relationship is forged and maintained by exploring the digital communication strategies employed by the third sector when interacting with communities. The results revealed that NGOs actively engage with communities through dialogical and networking approaches (Greenberg and MacAulay, 2009; Mergel, 2012). They also adopt a broadcasting approach through organisational posts that share insights into their daily operations and upcoming events. This transparency helps showcase their achievements and builds symbolic capital, bolstering their perceived legitimacy and effectiveness (Dupont, 2004). For instance, these posts might include information about successful community initiatives, partnerships with local organisations, or reports on crime reduction efforts, in an attempt to enhance their credibility and trustworthiness. Overall, these findings shed light on the approaches adopted by the third sector to try to cultivate and reinforce social ties with communities, facilitating their integration into the wider multi-agency partnership (Granovetter, 1973).

Although an essential aspect of this thesis, the strong social capital shared between the third sector and communities is not a novel concept. Previous research recognises the third sector as a unique entity that combines established external agencies and informal constituents. Scholars argue that NGOs possess structural and ambiguous organisational characteristics, granting them greater flexibility and adaptability than other external stakeholders, providing a "comparative advantage" for engaging community groups (Billis & Glennester, 1998). This is achieved by creating relatable and normative environments for community interactions to flourish in (Percy-Smith, 2003). This uniquely positions the third sector to establish and sustain self-supporting network ties with communities (Kelly, 2007).

### ***7.4.2. Utilising Informal Mechanisms of Collective Efficacy***

Consequently, non-profits can be viewed as hybrid stakeholders in a void between the formal and informal. As such they are crucial in promoting networking among and between communities (Blake et al., 2008) and directly fostering collective efficacy. NGOs leverage their unique position and alignment with communities to employ informal mechanisms of collective efficacy, akin to those employed by citizens. They achieve this using informal mechanisms, such as cohesion and trust-building exercises<sup>74</sup>, and employing informal social control while leveraging their organisational resources. This approach is akin to the informal mechanisms outlined by Sampson et al. (1997), and the third sector's hybrid position allows it to complement external interventions, such as community policing strategies, inclusive policies, and community infrastructure investment.

The proactive and reactive nature of the two informal mechanisms has been recognised as crucial factors in promoting collective efficacy. As previously mentioned, informal controls are reactive, addressing immediate issues for collective efficacy (Warner, 2014). Conversely, preventive measures like cohesion and mutual trust building can foster stronger bonds among residents, transcending prior perceived threats and strain (Sampson et al., 1997). These bonds integrate sub-communities through shared norms, values, and interests (Ferlander & Timms, 1999), serving as preventive means of upholding collective efficacy. While this thesis identifies this as a possible means of doing so, it makes no assumptions about its actual effect. This preventive approach aligns with the JSM and positive contact theory (Allport, 1954; Crandal & Eshleman, 2003), which suggest that positive interactions and accurate information can mitigate negative attitudes and reduce perceived threats.

Promoting community cohesion and trust is crucial for enhancing collective efficacy from a community cohesion perspective. Although heterogeneity may initially impact community cohesion negatively, studies indicate these effects can diminish over time

---

<sup>74</sup> It is important to note that although the existence of such exercises were explored their ultimate effectiveness was measured in this thesis.

(Twigg et al., 2010). Regular intergroup contact has increased outgroup trust and reduced perceived threats (McLaren, 2003), aligning with Allport's (1954) intergroup contact theory. Positive contact between different social groups can lead to social change by increasing knowledge and awareness of other cultures (Dovidio et al., 2017) and dispelling inaccurate information that fuels threat perceptions (Sternberg & Sternberg, 2008). Thus, cohesion and trust can be preventive measures to uphold collective efficacy.

Chapter Four's regression model suggests that NGOs prioritise cooperative and preventive measures rather than relying on reactive approaches associated with informal social controls. These organisations seek to address the underlying causes of tensions by prioritising knowledge and awareness campaigns as an attempted means to foster cohesion and mutual trust. This approach aligns with the objectives of the JSM in two ways. Firstly, by acting as a potential suppressive force against negative expressions, NGOs work towards establishing a shared understanding that such attitudes are unacceptable within the community. Secondly, by trying reducing justifications, constituents gain better knowledge and awareness of other groups through positive contacts facilitated by the awareness campaigns aimed at promoting cohesion and trust organised by NGOs (Crandall & Eshleman, 2003).

Chapter Six provides a comprehensive examination of the strategies utilised by NGOs to attempt to cultivate community cohesion and mutual trust to achieve collective efficacy. The strategies employed in communication today are diverse and encompass a wide range of approaches. According to Mergel's (2013) analysis, these approaches include communication strategies such as push, pull, networking, and transactional, as well as other sub-strategies that have been identified. In addition to these traditional methods, digital modalities such as hashtags, hyperlinks, and word count (Han et al., 2019) are also utilised in modern communication practices. Furthermore, dissemination, including information reach and depth, is an important aspect of communication methodologies (Van De Velde et al., 2015). Taken together, these diverse strategies and techniques have allowed for greater flexibility and effectiveness in communication practices across a variety of contexts. These approaches aim to facilitate knowledge exchange and promote

understanding, thereby enabling positive contact between and within communities, as postulated by the principles of positive contact theory (Allport, 1954).

A central facet of these strategies lies in their possible ability to mitigate negative opinions of sub-communities by disseminating accurate information. This information dissemination serves a dual purpose: to debunk negative stereotypes surrounding specific communities and to shed light on positive processes within these communities that may be unknown to others. The wide reach achieved through information dissemination can give way to achieving cohesion among diverse individuals within the community, encompassing various social categories that are instrumental in achieving collective efficacy (Van De Velde et al., 2015).

Concurrently, these organisations recognise the significance of information depth in addressing complex issues that demand contextual understanding. Providing in-depth portrayals using longer format inputs (usually through links) of community norms and engaging in meaningful conversations facilitates a comprehensive overview of diverse communities.

Chapter Six further found that NGOs leverage social media to gather public information to ensure accuracy and alignment with community beliefs. Interactive pull strategies engage citizens and utilise them as sensors for valuable insights (Mergel, 2012). By involving followers in activities, focus groups, town hall meetings, or surveys, organisations enhance social capital and gain a better understanding of community experiences and opinions (Meijer and Thaens, 2013). This participatory approach ensures that the disseminated information accurately represents the diverse perspectives within the communities served.

Furthermore, these approaches, guided by strong social ties with communities, facilitate positive contact experiences between and within communities. By promoting regular intergroup contact and facilitating accurate information exchange, NGOs could create opportunities for individuals, cultures, and ideas to engage in meaningful ways to build relationships based on mutual trust and respect (Sampson et al., 1997). These positive contact experiences encompass literal interactions between social actors and exposure to positive information related to these communities, which could contribute to greater mutual trust and understanding



(Dovidio et al., 2017). Positive contacts break down barriers, dispel stereotypes, and foster empathy among individuals from different social groups (Allport, 1954).

Through their knowledge and awareness campaigns, NGOs try to act as suppressive forces against causes of community tensions, establishing a shared understanding that such attitudes are unacceptable within the community. These campaigns could also serve as tools to reduce justifications by providing constituents with accurate information and promoting knowledge about other social groups (Crandall & Eshleman, 2003).

The third sector's approach was not confined to just the day-to-day but was also evident during and in the aftermath of specific events. In Chapter Five, the study revealed that attempted mutual trust and cohesion-building processes were also observed alongside reactive informal social controls after community tension trigger events. This indicates that informal social control mechanisms focused on trust and cohesion were implemented in addition to reactive measures. NGOs played a key role by organising proactive welcoming events (Sutter, 2019) to try and foster online and offline collective efficacy. These efforts aimed to generate comprehensive support for refugees and facilitate positive interactions between local constituents and newcomers, thereby building mutual trust (Allport, 1954).

Although Chapter Four emphasises the importance of preventive informal mechanisms, specifically the attempted establishment of mutual trust, it is crucial to recognise that NGOs also actively participate in informal social controls. This finding is consistent with scholarly perspectives highlighting the interconnected nature of informal mechanisms, suggesting that building mutual trust and engaging in informal social controls are mutually reinforcing (Simmons et al., 2005). Creating an environment conducive to community engagement and support for public interventions requires cohesion-building efforts and addressing tensions face-on. Therefore, while NGOs primarily focus on promoting cohesion and deterring, they could also assume a critical role in responding to and condemning negative behaviours and tensions as they arise. As discussed earlier, informal social control, characterised as a reactive mechanism in the pursuit of collective efficacy (Warner, 2014), involves a range of actions citizens take to address unacceptable behaviour within their community (Groff, 2015, p. 90). Within the framework of collective

efficacy, these actions include citizen interventions aimed at advancing the common good (Sampson, 2012), specifically for community tensions; a key aspect involves the proactive engagement with negative attitudes to confront and address them.

Chapter Five provided evidence of NGOs engaging in informal social controls. Notably, these groups were integrated into pro-refugee activist networks, blurring the lines between the third sector and community members, and emphasising their hybrid nature. The study found that NGOs played a crucial role in promoting collective efficacy by challenging and discrediting statements based on false premises and discriminatory generalisations irrelevant to the actual asylum accommodation. Through such interventions, they effectively debunked labels such as "illegal," "terrorists," or "rapists" that were unfairly attributed to refugees, signalling their rejection as unacceptable and incorrect to the wider community (Groff, 2015). This reflects the debunking processes inherent in their day-to-day attempted trust-building approaches, further highlighting the interconnectedness of informal mechanisms (Simmons et al., 2005).

These interventions serve as a force to suppress antagonistic attitudes and behaviours among both present and potential offenders (Crandall & Eshleman, 2005: 245). The impact of these interventions on collective efficacy can be measured through engagement metrics on Twitter, which demonstrated strong public endorsement for third-sector interventions, reflecting a high degree of collective efficacy and robust social ties between communities and the third sector. These cases can be understood as direct informal social controls. Additionally, in Chapter Five, NGOs were found to leverage external controls, such as on-platform reporting functions or reporting incidents to the police, mobilising external stakeholders to respond to antagonistic speech, labelled as indirect social controls (Sampson et al., 1997). This demonstrates that NGOs also utilise its close relations with inner-circle stakeholders (as evidenced in Chapter Four) of the CDN to respond and intervene during tensions by alerting them to incidents.

Chapter Six also examined the employment of informal social controls by NGOs. "Intervention posts," involving direct challenges to divisive attitudes, were associated with extremely positive responses across all engagement models, indicating strong

public endorsement when organisations intervene directly. This suggests that while these organisations primarily focus on trying to build cohesion in their day-to-day operations, the highest level of endorsement is achieved when they employ informal social controls to directly address divisive threats to cohesion. The general cultivation of community understanding and trust may bolster the widespread endorsement of such interventions. As Simmons et al. (2005) argue, informal mechanisms are not mutually exclusive. Establishing environments characterised by mutual trust and cohesion is critical for fostering a context where communities are more likely to participate in or support public interventions. This is particularly true when organisations encourage communities to mobilise collective responses, such as public protests (Delli Carpini et al., 2004), as seen in Chapter Six.

### **7.4.3. Summary**

This thesis provides compelling evidence that the third sector can undertake informal mechanisms of collective efficacy, previously attributed solely to informal actors within communities (Sampson et al., 1997). Nevertheless, it is important to note that these groups assume a hybrid role, wherein they can operate in both informal and formal capacities. This assertion is supported by Chapters Four and Six findings, illustrating their role as gateways (Legg, 2021) for community engagement within the CDN. In addition to their role as facilitators of cooperation, novel insights from Chapter Six reveal that the third sector can try to contribute to cohesion through more formal means. Specifically, they offer transactional environments to their constituents (Mergel, 2012). Notably, this involves making appeals for donations, which can be utilised to benefit minority groups through various channels. The study observed donation appeals manifested in diverse ways, including prize giveaways, trivia contests, and product sales, by previous research on transactional incentives employed by non-profit organisations (Castillo et al., 2014).

Significantly, one prominent finding highlights the extent to which the third sector provides support services for minority groups. This aligns with prior research on transactional strategies in the public sector, where there is an increasing demand for online customer service experiences (Poister and Thomas, 2007). In the case of the

third sector, services encompass a wide range of forms, such as information packs and hotlines for newly arrived migrants and victim support for individuals affected by hate crimes, hate speech, and domestic violence. The study's findings indicate that these support services are exceptionally well received, underscoring the third sector's invaluable role as a partner in assisting vulnerable communities. As a result, it is imperative to further integrate the third sector into existing systems to enhance the provision of support moving forward.

### **7.5. Cohesion Delivery in Digital Society**

Reflecting on the central question of this thesis (see Chapter One) one of the most significant dynamics shaping the contemporary Western world is the emergence of the digital society (Powell et al., 2018). The advent of digital society has led to a profound restructuring of social life, with digital technologies and communications playing a central role (Wildemeersch & Jütte, 2017). This transformative shift encompasses various spheres, including but not limited to the economic and social realms. It is crucial to recognise that the emergence of digital society has influenced how community tensions can arise and how they are addressed to pursue cohesion.

Therefore, a fundamental theme explored in this thesis is the impact of digital society on cohesion in Wales. Within this context, several sub-themes were carefully examined. These sub-themes encompassed the novel threats to cohesion in digital society, the adaptive nature of inter-agency cooperation facilitated by digital technologies, and how stakeholders can effectively employ digital technologies to mitigate tensions, and promote cohesion. This section aims to delve into these considerations, synthesising the findings from each empirical chapter and contextualising them in existing scholarly work within the conceptual framework of digital society. By doing so, a comprehensive understanding of the multifaceted impacts of digital society on community cohesion in Wales can be achieved.

### ***7.5.1. Emerging Threats to Cohesion in Digital Society***

Examining the new threats to cohesion associated with digital technologies (Hawdon et al., 2015) is crucial before delving into how cohesion delivery has adapted in the digital age. This was explored in Chapter Five, where a case study involving a trigger event was employed to investigate the online manifestation of anti-refugee sentiments during heightened tensions. A notable finding was the positive correlation between an offline event and its rise in digital tensions. While these findings are not novel, as previous studies have already established this relationship (see Williams & Burnap, 2016; Oskanen et al., 2020; Ozalp et al., 2020; Czymara et al., 2022), this thesis further contributes by demonstrating that this association can exist in the aftermath of localised cases of community tensions, such as in a small Welsh town. This expands the understanding that the association between offline tensions and online ones may be stronger than initially anticipated and can occur in a wider range of social contexts, not just following nationally significant events like terrorist attacks in London that dominate news cycles (Williams and Burnap, 2016).

Besides the connection between online and offline tensions, the study also identified several emerging threats to cohesion specifically generated by digital society. For instance, the impact of social media on localised tensions, like those observed in the Penally asylum accommodation, extends beyond the immediate geographical area, reaching other parts of the UK and taking on an exogenous effect (Williams, 2021). This reflects the concept of the network society proposed by Castells (2009), demonstrating how local issues can transcend spatial and temporal boundaries, enabling individuals with minimal involvement to engage actively by consuming information and participating in online discussions.

Another concerning discovery revolves around the heightened probability of antagonistic speech occurring with greater frequency and volatility when perpetrated by users utilising anonymous accounts and digital aliases. Perpetrators, shielded by anonymity, spread antagonistic messages using unfounded and factually incorrect statements, such as labelling refugees as "terrorists" or "rapists." According to Jaishankar's (2008) space transition theory, social media can facilitate the migration tensions from offline to digital spaces, amplifying the expression of discriminatory

feelings, particularly when individuals can fully anonymise themselves through identity flexibility and dissociative anonymity. This aligns with Festinger's (1952) deindividuation theory, which suggests that individuals are more prone to engaging in deviant behaviour when they believe they cannot be personally identified (Festinger, 1952). This process of disinhibition empowers individuals to offend or commit more serious offences than they would otherwise (Joinson, 1998). These explanations resonate with Crandall and Eshleman's (2003) JSM, which suggests that suppression primarily upholds a positive social identity aligned with social norms. However, perceived anonymity diminishes the significance of existing suppressions as individuals can easily detach their words from their identity.

The concept of space transition theory has been expanded by the findings of this study, which identified additional factors that contribute to community tensions facilitated by social media. These factors are unique to social media platforms and do not exist in offline capacities. For example, the study found that on-platform modalities such as hashtags and hyperlinks play distinct roles. Hashtags were found to have a positive association with anti-refugee content, corroborating previous research (Williams and Burnao, 2016). This is a cause for concern, as evidence links these elements to an increased spread through enhanced visibility in search results (Zappavigna, 2011). This suggests that they could potentially amplify and normalise negative attitudes. On the other hand, hyperlinks were negatively associated, aligning with previous research (Williams and Burnap, 2016: 227) that suggested divisive and extreme tweets are less likely to include hyperlinks to content that contradicts racist opinions or biased rumours. Hyperlink analysis supported this notion, as many directed users to traditional institutions like news outlets (e.g., BBC articles) or police press statements related to the camps, which were less predictive of anti-refugee content in the models.

Additionally, a positive correlation was discovered between the use of threatening language and anti-refugee content. This finding is consistent with prior research demonstrating that threatening language often accompanies expressions of animosity (Roxell, 2011). Consequently, this builds upon the evidence of a relationship between offline and online tensions, indicating that digital exchanges not only respond to offline events but may also contain threatening language that fuels

further tensions. This escalation of tensions can occur online and offline, creating a cyclical effect where each reinforces the other, amplifying the overall level of tension in both spaces.

Jaishankar (2008) emphasises the significant factor of deterrence, or lack thereof, about cyberhate. Scholars acknowledge cyberhate's diverse and ever-changing nature (Brown, 2018), presenting challenges to policing and legislative processes (Bakalis, 2018). Consequently, digital spaces exhibit minimal deterrence compared to offline environments, as the associated suppressive forces of being caught and identified are also minimal, especially when much of the dialogue is more antagonistic than “hateful” per se. This becomes particularly evident during trigger events when internet traffic surges, thereby straining the response system.

Chapter Five provides evidence of unmoderated content on digital platforms. First, it revealed that closed network platforms like Facebook groups were frequently used to propagate anti-refugee sentiments. This trend raises concerns about the discreet operations of such movements within hidden underbellies that are harder to moderate and report as they facilitate more private communications. Closed network spaces amplify radical views within echo chambers of like-minded individuals in the network society, linked by shared intolerant perceptions (Castells, 2009). Literature suggests that closed networks, including Facebook groups, can amplify polarisation potentially leading to real-world coordinated incidents (Harel et al., 2020). While closed network structures are commonly associated with such phenomena, Chapter Five's findings also indicate that open-network platforms like Twitter can facilitate the development of unmoderated content. The data from Chapter Five was collected retrospectively, indicating that social media moderators did not remove or moderate the extreme content found on Twitter, even a significant amount of time after publication. This finding is particularly concerning, considering that it predates Elon Musk's Twitter takeover, during which higher-level censorship existed (Curwen, 2022). Moreover, Musk's vocal criticism of censorship and moderation practices could embolden users to engage in more extreme content, knowing that there may be less likelihood of repercussions. This could exacerbate existing challenges faced by marginalised communities, who are disproportionately targeted by online harassment and discrimination. As we navigate the complexities of online discourse

under Musk's stewardship of Twitter, it becomes increasingly crucial to remain vigilant and advocate for policies that prioritise the well-being of all users. Efforts to diffuse tensions must not waver in the face of changing leadership or ideological shifts. Instead, they should be reinforced with a renewed commitment to fostering a digital space where respectful dialogue thrives, and individuals feel safe to express themselves without fear of harassment or intimidation.

As mentioned, digital society has enabled discriminatory attitudes to thrive in social media discussion networks, leading to more concerted and organised antagonistic expressions that create tensions. Chapter Five's findings highlight core-periphery structures within these networks, forming echo chambers where like-minded individuals amplify and encourage each other's content. This aligns with theories of group socialisation and contagion effects as antagonistic sentiments become normalised and spread among individuals (Sunstein, 2017). Social media platforms like Twitter act as polarisation amplifiers, galvanising and propagating negative narratives targeting outgroups (Ellis, 2002; Daniels, 2008; Sunstein, 2017). The study's analysis of discussion networks related to anti-refugee attitudes supports the notion of echo chambers, with central nodes representing the lead propagators and mentions playing a significant role in promoting these (Ozalp et al., 2020). However, it is important to recognise that these closed echo chambers can result in the development of more radical perspectives among those propagating divisive content. While confined to a closed-off group, these individuals may become increasingly extreme in their views. However, as echo chambers become more radical, they became further disconnected from public perceptions and received limited public endorsement, therefore only radicalising a select few within them (as seen in Ozalp et al., 2020).

### ***7.5.2. Digital Technologies to Enhance Nodal Governance***

With the increasing challenges the digital society poses, it is crucial to explore how digital technologies can be utilised to foster cohesion instead. One way explored is how these technologies can be leveraged to improve coordination among stakeholders involved in delivering cohesion. This coordination has been observed to



occur among formal partners and on social media platforms, where it helps establish and strengthen social capital within communities. Chapter Four sheds light on the potential impact of the COVID-19 pandemic on the collaborative efforts of the multiagency network responsible for cohesion delivery. This impact can be attributed, in part, to the greater reliance on digital communication methods, particularly the increasing popularity of real-time communication systems (RTCs).

Findings in Chapter Four highlight that the COVID-19 pandemic has increased inter-agency cooperation despite initial concerns. Previous research has demonstrated the significant role of RTC platforms, such as Microsoft Teams, in facilitating virtual meetings and coordination among organisations (Gray et al., 2020; Shakespeare et al., 2022). Building upon this existing body of work, this thesis further underscores the transformative potential of RTCs in enhancing inter-agency cooperation, specifically within CDN and initiatives. As such, RTCs have effectively improved engagement, including partner liaison, but also saved travel time, leading to increased productivity and stakeholder interaction. However, the lack of a standardised RTC platform has posed challenges, with discrepancies in platform use.

Nevertheless, the thesis findings indicate that organisations have overcome these challenges and worked closely together during the pandemic. Virtual meetings have enhanced communication, information exchange, and engagement with partners. Moreover, virtual meetings have promoted equality and cooperation among agencies and regions, highlighting a less Cardiff-centric approach. The increased reliance on RTCs, as emphasised in the thesis, has prompted organisations to recognise their potential to enhance internal and external communication and cooperation, driving a broader digital transformation of workplaces beyond the pandemic (Savic, 2020).

### ***7.5.3. Curtailing Community Tensions During Trigger Events***

Chapter Five presented empirical evidence showcasing the effectiveness of digital spaces, specifically social media, in promoting cohesion during heightened tensions following trigger events. By examining informal communities and specialised

cohesion delivery stakeholders, this chapter provides a comprehensive overview of the practical strategies employed in the aim of achieving digital cohesion. The findings underscore the significant role social media could play as a catalyst for fostering cohesion in the digital realm. By harnessing social media, individuals and organisations can navigate the complexities of heightened tensions to cultivate a sense of togetherness and unity. This provides a valuable glimpse into the practical implementation of strategies that could contribute to digital cohesion. This thesis analysed online interactions to understand how they are specifically used to diffuse heightened tensions. Platform engagement metrics and qualitative extracts demonstrate the different approaches endorsed or rejected. Consistent with previous research (Williams and Burnap, 2016; Ozalp et al., 2020), results showed that retweets do not perpetuate antagonistic speech, indicating limited endorsement among the general public. This finding suggests the existence of digital collective efficacy, where digital constituents come together to reject discriminatory beliefs and foster a sense of community and solidarity (Sampson et al., 1997).

Additionally, the study emphasises the positive association between pro-refugee activists, comprising individuals and NGOs, and the dissemination of information on social media platforms. Pro-refugee activists leverage social media to amplify their message, garner support through retweets, likes, and replies, and spread positive narratives. By mobilising communities and facilitating open dialogue, social media could enable the collective voice of individuals to drown out discriminatory ideologies (Ozalp et al., 2020). The transformative impact of social media extends beyond specific activist movements. Established institutions such as police, media outlets, and political figures are crucial in shaping public opinion in digital spaces. Results showed that these entities were all negatively associated with publishing anti-refugee content. Leveraging their symbolic capital (Dupont, 2004), they have the capacity and potential to disseminate accurate and unbiased information, challenging misinformation, and debunking stereotypes.

Furthermore, the study explores the role of the police in utilising social media platforms, showing their active engagement in open discussions with the public. While they receive fewer endorsements through measures such as likes and retweets, their primary mode of operation involves replies that allow for transparency

and collaboration. This enabled the police to communicate accurate and up-to-the-minute information about such tensions to communities. Such use of social media platforms acts as bridges, breaking down barriers between law enforcement and the community, potentially promoting trust, cooperation, and practical problem-solving, which contributes to building more positive images (Mawby, 2002). This indicates that social media could play a transformative role in enhancing the police's trustworthiness and symbolic capital (Dupont, 2004), if harnessed in the correct way allowing them to effectively combat tensions.

The findings reflect the media's<sup>75</sup> continued influence in shaping public discourse. While the digital age has transformed information dissemination, media outlets still possess significant reach and impact, supporting the classic criminological notion that conventional media plays a role in shaping public discourse following important events by "setting the agenda" and "transmitting the images" (Cohen, 2002). The study shows that the media and press headlines positively affect engagement metrics. This indicates that responsible and unbiased reporting by the media is crucial and could play a vital role in challenging negative attitudes by setting the agenda, transmitting accurate information, and debunking false claims (Cohen, 2002; Williams and Burnap, 2016). However, it is also a cautionary finding, demonstrating that the media can still influence public perceptions in the digital age. If they were to report negatively or propagate discriminatory content, they could catalyse further tensions, as seen in similar studies (King & Sutton, 2014; Moore and Ramsey, 2017; Sadique et al., 2018). Therefore, although this particular example shows the potential positive use of influence to diffuse negative sentiments towards refugees, the continued influence of the media, even in the digital society, should be regarded with caution.

---

<sup>75</sup> The media was predominately measured by exploring written news headlines and readers should note other media forms could exhibit notable differences.

#### ***7.5.4. Daily Strategies Against Digital Tensions and Divisive Attitudes***

While Chapter Five provided evidence on how organisations can utilise social media during heightened tensions, Chapter Six highlighted the strategies used to try to promote cohesion social media in everyday situations. Currently, the only stakeholders with a digital presence solely focused on promoting cohesion are NGOs. Therefore, the third sector's utilisation of social media in Chapter Six can be seen as a potential blueprint for other organisations to follow. In this regard, government bodies or the police, for instance, could establish social media accounts dedicated explicitly to cohesion and equality. By adopting and implementing successful approaches identified from the third sector, these entities can leverage social media to promote inclusivity and mitigate tensions on their main accounts and through potential sub-accounts.

In the context of digital society, the definition of community has evolved. Castells (1997) argues that digitalising social life has led to communities no longer being organised solely based on spatial or temporal dimensions. Instead, many communities now exist in digital spaces and differ from their offline counterparts in size and composition (Dutton, 1996). Nevertheless, scholars agree that, like offline contexts, digital communities also need to establish norms of behaviour and restrict online transgressions (Postmes et al., 1998). This helps cultivate a "netiquette" (Shea, 1994) wherein digital constituents mutually understand that these behaviours are unacceptable, and communities are more likely to intervene and signal when such behaviour occurs.

Chapter Six delves into organisations' regular utilisation of social media, focusing on establishing digital cohesion through implementing various strategies (as outlined in Mergel's (2012) framework presented in the same chapter). Each approach serves a distinct purpose while sometimes also interconnected with one another. Firstly, knowledge and awareness campaigns are employed trying to educate and inform the public about issues related to minority rights and cultures. This also encompasses sharing information and resources to raise awareness and promote understanding of the impact of victimisation on individuals and communities.

Additionally, interventions are employed in attempt to address unacceptable behaviour or actions, and organising community events and protests that promote inclusivity. Both approaches, integral components of the overarching strategy, significantly contribute to developing collective efficacy (Sampson et al., 1997), enabling digital communities to actively address and counteract divisive attitudes. As cohesion develops within the community, constituents become empowered to challenge divisive attitudes, fostering a dynamic synergy where collective efficacy is significantly enhanced, as Simmons et al. (2005) emphasised. This reciprocal process sees the empowerment of individuals to mobilise against reinforcing the overall sense of cohesion and creating a positive feedback loop. In this context, the community could actively reject intolerant views, explicitly signalling that such perspectives are unacceptable and should not be normalised. The rejection of intolerance could therefore become a shared value within the community, further strengthening the cohesion and trust among its members. This interconnectedness, or synergy, has previously been found to amplify the impact of both strategies, working in tandem to cultivate netiquette (Shea, 1994) and establishing a digital environment where hate is actively challenged and not allowed to become the norm. It is important to note that while such strategies outlined in this thesis aim at promoting cohesion, no assumptions are made about their ultimate effectiveness, a compelling topic for future research.

Building upon these principles, another pivotal approach involves the implementation of dialogical approaches designed to enhance social capital within informal digital communities and with formal stakeholders on social media platforms. By cultivating robust social capital through these dialogues, organisations could strengthen the bonds within the digital community and could aid in creating a collaborative environment that directly reinforces the objectives of building cohesion and trust. This thesis, however, did not, directly measure the effectiveness of this and cannot make any claims about the overall impact. If social capital grows through these dialogical engagements, it could become a driving force behind the collective efficacy necessary for community building. The interactions, whether occurring online or extending to offline events, could serve as a dynamic platform for community members and stakeholders to strengthen social ties and exchange insights and experiences.

In this context, it allows for an information sensors concept (Mergel, 2013), with organisations gaining valuable insights from community members and practitioners. This dual perspective, drawn from lived experiences and professional knowledge (what Dupont (2004) refers to as cultural capital), not only refines the organisation's awareness of issues but also guides the development of more effective and informed strategies. The synergy between dialogical approaches enhanced social capital, and inclusive information gathering becomes a linchpin for organisations striving to foster cohesive digital communities, instil trust, and directly. This interconnected strategy ensures that the collective efforts are grounded in a comprehensive understanding of diverse perspectives, ultimately leading to more impactful and sustainable outcomes.

Additionally, they leverage social media to offer support services and solicit donations. Support services, in particular, were particularly effective. Consequently, harnessing support services on social media holds significant potential for expansion, not only within the third sector but also through strategic collaborations with other partners. Scholars consistently emphasise the importance of providing support services to marginalised individuals. However, concerns persist about the limited accessibility of these services (Hardy and Chakraborti, 2020). They highlighted that only a privileged few, often those in more economically or socially stable circumstances, know such services. Notably, even among these individuals, a significant portion remains uninformed about these resources. The research findings suggest significant engagement with support service approaches, indicating the potential of social media, in addressing this issue by enhancing the visibility and accessibility of these services to a broader audience. Acknowledging that not everyone has access to social media, potentially excluding some individuals is important. Nevertheless, raising awareness through these platforms could provide advantages in reaching a larger and more diverse population.

The findings demonstrated that social media platforms, such as Twitter, could potentially have a noteworthy impact in fostering cohesion and mutual trust, even in everyday situations rather than solely during the aftermaths of trigger events. Consequently, these approaches serve as evidence of the transformative capabilities of social media in promoting social inclusion and facilitating active community participation (Vance et al., 2009). In light of these insights, other agencies must

recognise and adopt these effective approaches on social media, both on their main accounts and by developing specific inclusion and equality sub-accounts. By doing so, they can promote cohesion, mitigating tensions, and foster inclusivity, thereby creating positive change in digital and offline spaces.

#### ***7.5.5. Digital Tools for Tension Monitoring***

A demand for a standardised dashboard tool capable of rapidly and efficiently capturing and analysing social media data was highlighted in Chapter Four. A significant majority (82.5%) of respondents believed that additional methods of online tension monitoring should be implemented within the network. All cluster groups exhibited substantial support, with the police and community coordinators demonstrating the highest levels of agreement (89.9% and 87.5%, respectively). Stakeholders expressed dissatisfaction with current manual monitoring processes, citing their time-consuming and inefficient nature. They argued that a tension monitoring dashboard could considerably reduce the resources and time invested in manual monitoring.

Moreover, stakeholders recognised social media as a valuable source for gathering data on tensions before they escalate, underscoring the preventive potential of such a tool. Furthermore, stakeholders identified limitations within the existing reactionary framework for documenting digital tensions, particularly within organisations like the police. The lack of reported tensions and communities' reluctance to report issues indicated a deficiency in symbolic capital and a breakdown in communication between groups. The open output revealed that to address these challenges, stakeholders advocated for a new tension monitoring tool that is “automated” and reliable to be both “time and cost-effective”, and capable of capturing all online tensions, including hidden ones in real-time. In addition the ability to geographically locate tension hotspots, the absence of biases towards specific demographics, user-friendly monitoring interfaces, and the provision of intelligence on resulting offline tensions. This approach would enable covert monitoring of genuine tensions and the identification of larger trends for effective prevention strategies. These insights

informed the development of potential tools for monitoring tensions, such as the Hate Lab dashboard.

The Hate Lab dashboard, which focuses on monitoring and analysing hate speech and online abuse that leads to wider tensions, was identified as a promising tool by an academic contributor. It offers access to a wider range of social media platforms than previous systems like CAT-D. The HateLab Platform's interactive and user-friendly interface allows for the swift aggregation and visualisation of complex data, facilitating the examination of hate speech trends over time and location.

Additionally, the platform's potential to identify patterns of online inter-group tension that correspond to offline phenomena and inform strategies for addressing them was emphasised. In summary, Chapter Four's findings emphasise strong support for implementing formal controls (Sampson, 2001), such as online tension monitoring methods and the necessity of a comprehensive monitoring tool. The Hate Lab dashboard emerged as a promising solution that fulfils the requirements highlighted by stakeholders, providing cost-effective and valuable monitoring capabilities for inter-group tensions in Wales. Furthermore, the value and practicality of potential monitoring tools were exemplified in Chapter Five by utilising an illustrative dashboard for data visualisation. While the researcher manually constructed this particular dashboard in this thesis, the underlying data and visualisation techniques it employed indicate emerging tools, such as HateLab or COSMOS, that hold significant promise and could have a meaningful impact. Consequently, this dashboard serves as a representative example to showcase the potential benefits such tools can offer practitioners despite the impracticality of manually implementing similar processes.



## Chapter 8

### Summary

---

#### 8.1. Introduction

This thesis investigated cohesion delivery in modern society, focusing on Wales as a case study. Three distinct objectives were developed, each guiding separate empirical studies that addressed research questions and hypotheses related to the central thesis. The first study explored cooperation patterns within the CDN in Wales, examining how social actors collaborated to maintain community cohesion during the COVID-19 pandemic and the role of social media in community tensions. The second study investigated the dynamics of community tensions and community responses following the establishment of Penally asylum accommodation in Wales. This study analysed the interaction of social actors and the impact of emerging digital technologies on the development of tensions, their propagation and collective responses. The final study examined the use of social media by NGOs to address social injustices and promote community cohesion. This study analysed the digital communication strategies employed by these organisations in an attempt to reduce tensions.

A pragmatist methodological approach was developed after reviewing the relevant literature to allow for an exploratory investigation of emerging concepts that have not been extensively explored in academia. This approach enabled traditional and new approaches to investigate the research questions and test hypotheses. For example, digital surveys, factor reduction and MDS techniques were used to collect data for multi-agency cooperation patterns (Levi and Williams, 2013), extensive locomotive online data was collected around trigger events, allowing for a fine-grained analysis of temporal phenomena (Edwards et al., 2013), and digital communication frameworks typologies were used for content analyses of communication data (Mergel, 2013).

The discussion section synthesised the key findings from the three empirical chapters, offering a comprehensive overview of cohesion delivery in the contemporary world.

By using Wales as a specific case study concerning related literature and conceptual frameworks, this research has shed light on the broad research question of "How is community cohesion maintained and promoted in contemporary Wales?". The investigation encompassed various considerations, including external and informal mechanisms of cohesion delivery, the unique role of the third sector in assuming a hybrid role in promoting cohesion, and the impact of the digital society on both posing new threats to cohesion and facilitating transformative effects in cohesion delivery. This study has contributed to a deeper understanding of the dynamics and complexities of sustaining community cohesion in modern-day Wales by addressing these multifaceted aspects. The insights gained from this research offer valuable implications for policymakers, practitioners, and researchers seeking to foster inclusive and resilient communities in today's rapidly evolving societal landscape. This chapter provides a contribution statement of key findings following the discussion section. Finally, the limitations of the research are outlined, along with potential future research directions.

## **8.2. Contribution Statement**

This statement highlights the value of the research conducted in this thesis, emphasising its significance and impact. Within this section, three distinct contributions are delineated: knowledge, methodological, and practical contributions applicable to both academia and practitioners (Bergh et al., 2022). The knowledge contributions aim to expand existing academic understanding by addressing critical gaps in the field. Pioneering methods employed within this thesis will be highlighted as methodological contributions, which enable valuable insights. Lastly, actionable findings and recommendations will be presented as practical contributions, bridging the gap between academia and real-world application.

### **8.2.1. Knowledge Contributions**

Before delving into the specific contributions across the three empirical chapters, it is valuable to consider the broader knowledge generated by this thesis. One significant contribution is that it represents the first research explicitly exploring cohesion delivery in Wales. Consequently, it offers a comprehensive overview of the specific strategies implemented in Wales. Despite this, many common themes, especially those observed in digital spaces that transcend geographical boundaries (Castells, 1997), are considered applicable beyond Wales and are universally important. Similarly, the interagency and community-based findings potentially reflect other populations outside Wales, particularly within Western cultures. Thus, while this research focuses on Wales, its findings extend beyond its geographical confines, providing valuable insights into threats to cohesion and effective strategies for addressing them in a more general context. Another significant contribution of this thesis is applying Sampson et al.'s (1997) concept of collective efficacy specifically to cohesion contexts. While previous studies (Costello et al., 2017; Ozalp et al., 2020) have touched upon this concept, this research builds upon them by offering a more comprehensive examination. It explores both the informal and formal components of collective efficacy in relation to cohesion, going beyond existing work by identifying specific partners, such as the third sector, that serve as hybrids between informal and formal approaches. These partners are able to employ informal strategies despite being formal entities, thus expanding on our understanding of the concept of collective efficacy in the context of cohesion.

Chapter Four presented the first-ever network analysis of the Welsh CDN, which is not a formally established network. The analysis provided valuable insights into the cooperation patterns among key stakeholder groups. One of the significant findings was that NGOs and cohesion officers play a crucial role in integrating communities and their perspectives into multi-agency responses. However, it was observed that many community groups remain unaware of the existence of cohesion teams. This highlights the untapped potential of specific roles, such as cohesion officers, who possess greater symbolic capital with communities. Consequently, communities tend to feel more comfortable discussing and reporting issues to these partners compared to more traditional bodies like the police. Interestingly, disability focus was focused

on other protected characteristics. Moreover, the study delved into the impact of COVID-19 on multi-agency partnerships, specifically in the context of cohesion delivery. The pandemic resulted in an increased use of digital technologies to enhance cooperation frequency, which has become a lasting legacy. The findings of this research have significant value and contribute to the broader knowledge in the field. The results were converted into a research article<sup>76</sup> published in the British Journal of Community Justice, granting the academic community wider access to these knowledge contributions.

Chapter Five investigated the dynamics of community tensions and responses in the aftermath of a localised trigger event - the establishment of Penally asylum accommodation in Wales. It examined how various social actors, including politicians, the media, civil society organisations, and individual citizens, interact with the propagation of anti-refugee attitudes and the attempted maintenance of community cohesion in times of crisis. This study is built on an emerging criminological discipline that analyses the temporal dimensions of tensions rather than spatial (King and Sutton, 2013). To date, the majority of temporal studies have focussed on offline contexts (Legewie, 2013; King and Sutton, 2014; Hanes and Machin, 2014). However, this study extended the investigation to the realm of digital tensions, drawing on recent research that has shown similar patterns in online contexts (Ozalp et al., 2020; Czymara et al., 2022). While the exploration of trigger events in digital spaces is not entirely novel, this research expanded on these pioneering studies by contributing to knowledge in several further ways, filling knowledge gaps. The first contribution is exploring an extremely localised trigger event in a town in Wales rather than large-scale events like Trump's election, the Brexit referendum, or national terrorist incidents (King and Sutton, 2014; Hanes and Machin, 2014; Williams and Burnap, 2016). It builds on the surrounding literature by investigating whether local events can cause similar temporal shocks and exogenous effects on information dissemination and community tensions. Evidence showed digital tensions and discourse spreading well beyond Penally and across the

---

<sup>76</sup> The research article is titled "Combating Community Tensions in Wales: Mapping the Cooperation Space for Multi-Agency Cohesion Delivery" and can be accessed at the following link: <https://doi.org/10.48411/Odtk-wg13>

UK as a whole, thereby contributing to the existing literature by providing the first examination of micro-trigger events where the event originates within a small local area.

The second contribution is further exploring the role that discussion networks and echo chambers play in the propagation of discriminatory attitudes after such events. Previous work identified a small number of individuals operating in a close-knit echo chamber responsible for the majority of antagonistic speech (Ozalp et al., 2020). This study expanded on such theorisations by building reply network visualisations that filtered the nodes and edges engaging in anti-refugee speech, providing visual evidence of a core-periphery structure (Borgatti & Everett, 2000). This visual evidence showed a small echo chamber of the main propagators supporting previous theorisations (Ozalp et al., 2020). Additionally, examining other engagement metrics beyond retweets, such as likes, revealed a previously unrealised aspect of passive support for these sentiments, indicating a dark underbelly of supporters who may not actively propagate them by retweeting but express their support through likes.

Another significant contribution is the exploration of social factors by differentiating between far-right political agents and far-right activists. Previous trigger event research treated far-right agents as one entity (Williams and Burnap, 2016), whereas this study split them, revealing staggering differences between the two groups. While both were still associated with anti-refugee sentiment, the findings showed that far-right activists were more prone to volatile sentiments. This novel finding was further conceptualised by considering Festinger's (1952) theory of deindividuation, where most activists used anonymous accounts and were more prone to publish more anti-refugee content. This finding demonstrates the potential for future work to further split social agents into subcategories for more accurate insights into their differing roles in the wake of such events.

Chapter Six involved a novel content analysis of how NGOs utilise social media platforms to promote cohesion on a daily basis. Previous studies (Williams and Burnap, 2016; Ozalp et al., 2020) have primarily focused on how cohesion stakeholders respond to specific trigger events, as discussed in Chapter Five. These

reactive approaches aim to counter divisive attitudes and can be considered social controls, as outlined by Sampson et al. (1997). This chapter builds upon existing research by exploring proactive strategies to prevent tensions and promote cohesion through daily social media activities. By examining digital approaches to fostering mutual understanding and cohesion, this study provides a more comprehensive exploration of Sampson et al.'s collective efficacy (1997), expanding the scope beyond reactive measures. To achieve this, the study adopts Mergel's (2013) framework, which was previously applied solely in the context of public sector organisations. By employing this framework to analyse NGOs with a minority right focus, this study contributes to the existing literature on the role of the third sector in promoting collective efficacy. Consequently, it demonstrates the versatility of Mergel's framework in various contexts related to digital communications, extending beyond the boundaries of the public sector. The findings highlight the instrumental role of the third sector and social media in promoting cohesion through a diverse range of strategies that can also be adopted by other stakeholders.

### ***8.2.2. Methodological contributions***

As Bergh et al. (2022) suggest, methodological contributions are challenging to define with precision. However, broadly speaking, they refer to the advancements and improvements in research methods within a given discipline, enabling scholars to conduct more accurate and reliable research. These contributions may include new techniques, tools, and approaches, as well as modifications to existing methods. By developing and implementing these innovations, researchers can address limitations in the current research techniques and improve the quality of their findings.

The first methodological contribution came in Chapter Four when implementing MDS for the analysis of multi-agency partnerships. MDS, widely recognised and used across diverse fields, efficiently uncovers patterns through the visual representation of quantitative proximity data.. MDS served as a tool to visualize communication patterns among cluster groups, building upon the foundational research of Levi and Williams (2013). To facilitate interpretation, additional analytical methods were

integrated, including PCA, regression, and qualitative data. MDS simplifies complex data, allowing researchers to easily interpret and visualize cooperation patterns (Coxton, 1982). The practical significance of employing MDS lies in its ability to simplify intricate data and enhance accessibility, particularly in research involving practitioners. This approach offers easily understandable visualizations for non-academic communities. The study further extended the application of MDS to visualize stakeholder focus on the nine protected characteristics. This innovative use revealed clusterings and separations, offering valuable insights into the attention given to specific characteristics. Consequently, this advancement not only broadens the utilisation of MDS for exploring network cooperation patterns but also delves into operational priorities within multi-agency partnerships, building upon the earlier work of Levi and Williams (2013).

Chapter Five expanded on existing work by examining digital tensions in relation to trigger events. This approach provided valuable insights into the online manifestation of anti-refugee attitudes following specific events and introduced methodological advancements through the utilisation of extensive locomotive approaches. These approaches allowed for a more comprehensive understanding of anti-refugee content propagation and its associated challenges. As noted by Edwards et al. (2013), extensive locomotive data enables a fine-grained analysis down to the second, providing a precise examination of phenomena. One noteworthy benefit of employing extensive locomotive approaches employed in this chapter is their ability to establish realistic perceptions and attitudes in real or near-real time, particularly when traditional methods like interviews or surveys are unlikely to engage perpetrators (Williams and Sloan, 2015). By adopting extensive locomotive approaches, the study minimised the risk of bias and achieved a more reliable means of capturing and understanding anti-refugee content (Drouhot et al., 2023).

While previous studies have applied extensive locomotive approaches (Williams and Burnap, 2016), this study further advanced the field by utilising such approaches over more gradual and extended time periods. Unlike prior research that predominantly focused on nationwide trigger events, this analysis expanded the scope by considering local issues related to the Penally asylum accommodation. This extension demonstrated the feasibility of implementing detailed temporal

analyses over significantly longer durations than previously explored. In conducting the study, big data emerged as the primary source among the four utilised, effectively complementing traditional data sources and broadening the research designs typically employed in extensive locomotive contexts (Morse, 2016; Snelson, 2016). By integrating these additional sources with big data, the study facilitated a more comprehensive exploration of tensions and propagation of divisive content, as well as the challenges associated with them.

### ***8.2.3. Practical contributions***

In addition to advancing theoretical knowledge and methodological innovations, research has the potential to make significant practical contributions and improve the functions of associated practitioners, as highlighted by Bartunek and Rynes (2010). In the context of this PhD thesis, as the researcher, I have actively engaged with the CDN, participating in multiple cohesion coordinator meetings, and serving as a member of the Hate and Community Tension Board (Cymru). Furthermore, the PhD received funding from the Welsh Government and Police. Through close collaboration and continuous communication with these agencies, I have ensured that the key findings and research outcomes are effectively shared, facilitating practical contributions during and after the studies. To streamline dissemination, a comprehensive summary of the research, accompanied by a concise executive summary and key findings, was created, and circulated among the board members and relevant stakeholders. This summary encapsulated the essence of the research and included recommendations specifically tailored for practitioners involved in all three empirical chapters. By sharing this information with practitioners, the research strives to bridge the gap between theoretical knowledge and practical application, fostering an environment of evidence-based decision-making and enhancing the effectiveness of associated professionals in their respective fields.



### 8.3. Limitations and Future Research

While this research has made contributions to our understanding of contemporary cohesion delivery and its real-world practical implications, as well as introducing innovative methodologies, it is important to acknowledge certain limitations that should be considered. Before addressing the specific limitations associated with each empirical chapter, it is important to outline the broader limitations of this thesis' inquiry. Firstly, it should be noted that all analyses were conducted solely on data from Wales. While many of the findings may have broader applicability to contemporary societies, it is crucial to acknowledge that some of the findings may be specific to Wales and may carry less significance when extrapolated to other regions. For instance, the specific cohesion roles described in this thesis are unique to Wales, and although they could serve as a blueprint for other nations, these roles may not exist elsewhere. Future research could explore the transferability of these roles, among other Wales-specific components, or investigate the existence of similar ones in different contexts. The second broad limitation pertains to the extensive scope of the concept of cohesion delivery. This thesis has provided valuable initial insights into several processes undertaken to mitigate tensions and promote cohesion. However, it is important to recognise that the thesis only scratches the surface, leaving many facets of cohesion delivery unexplored. Future research could delve into various aspects, such as examining the role of social media companies in censoring and restricting antagonistic speech, evaluating the effectiveness of specific governmental policies in fostering cohesion or investigating the contributions of other bodies, such as the private sector, towards achieving cohesion, among other potential areas of inquiry.

Chapter Four presented initial evidence of patterns in the CDN; however, it is crucial to consider certain limitations in the research design when interpreting the results. Upon reflection, three key limitations have been identified. Firstly, the use of a centralised Welsh Government stakeholder group is a limitation that needs to be acknowledged. While this was determined by Welsh Government policy to ensure a unified response, it is important to recognise that the Welsh Government itself is a multi-layered and complex entity with various factions, even within the community justice and equality branches. By only including one unified group representing the

Welsh Government, the research may overlook significant variations in cooperation experiences among different factions. This limitation extends to other respondents as well, where clustering together stakeholders like the Police, Welsh Government, community groups, and NGOs provided a macro-level analysis of inter-agency cooperation but lacked insights into meso-level and specialist roles, as well as micro-level perspectives within each agency. Future studies that delve into specific perceptions of different actors in various positions within these agencies would contribute to a more comprehensive understanding of the complete network. Although this research offered valuable insights as the first study of the network in any capacity, it is essential to address the second limitation associated with the use of survey-based open questions. The restricted nature of open questions limits the researcher's ability to explore network dynamics further. Follow-up studies with longer qualitative elements could adopt a mixed methods approach, allowing for a more in-depth exploration of the network dynamics. For example, conducting focus group exercises involving members from all stakeholder groups could prove effective in examining inter-agency relationships. The final limitation pertains to the use of non-metric perceptions in the cooperation Likert-scale questions. It is important to acknowledge that responder demographics, agency norms, and differing opinions on workplace dynamics and partnerships can significantly influence participants' responses on the scales. Future research could focus on studying these biases, shedding light on additional dynamics within the network.

Chapters Five and Six concentrated on social media data, and as a result, certain limitations associated with social media data emerged in these chapters. Both studies exclusively explored social media data derived from Twitter. While the methodology (further details provided in Appendix 4) justifies the use of Twitter as the most suitable platform for this thesis, relying solely on Twitter poses constraints on obtaining a more comprehensive understanding of the studied phenomenon in digital spaces. Platforms like Facebook, Instagram, and Snapchat offer content formats and features that are not present on Twitter. These platforms emphasise audio-visual content and allow for longer or unrestricted text, as noted by Han et al. (2019). Additionally, other platforms could provide valuable insights into different demographic groups that are not easily accessible through Twitter. This is particularly relevant when considering significant actors who propagate antagonistic

speech, as they may predominantly reside on platforms such as 4chan. Therefore, future research aiming to investigate cohesion in digital spaces should consider using novel social media data from platforms other than Twitter. Analysing data from multiple platforms will offer a more comprehensive understanding of digital activities related to cohesion.

Chapter Five of this thesis had several identified limitations that should be taken into consideration. Firstly, it was the only analysis within this thesis that examined threats to cohesion in digital spaces. However, the findings were based on a specific case study related to a policy change in a local area trigger event, which resulted in increased anti-refugee attitudes. Consequently, these findings provide limited insights into the development of day-to-day tensions unrelated to trigger events or trigger events related to different causes concerning different social groups.

Therefore, caution should be exercised when generalising the findings to other cases (Tsang, 2014). Future studies could address this limitation by exploring multiple case studies to identify common themes across different contexts. Another limitation pertains to the inclusion of offline hate incident data in relation to digital tensions.

The analysis only incorporated aggregated monthly offline hate incident data, and it was not possible to examine these on a daily basis. This limitation may be attributed to the localised nature of the case study, where analysing hate incidents on a daily basis may not yield significant results. However, future studies, especially on a larger scale, such as nationwide elections, should strive to collect and analyse hate incident data on a daily basis to ensure more accurate and comprehensive findings. While this study expanded on previous research by examining engagement metrics beyond retweets (including likes and replies), it is important to acknowledge that other engagement metrics, such as views and link clicks, could be measured to gain a deeper understanding of how information is disseminated within digital spaces. Incorporating these additional metrics in future research would provide a more comprehensive analysis.

The accuracy of the Perspective API classifier was evaluated using the F1 coefficient, which combines precision and recall values, and a threshold of  $>0.70$  was considered acceptable (van Rijsbergen, 1979). The Perspective API for identity attacks surpassed this threshold, indicating its suitability as a dependent variable for

anti-refugee content. This demonstrates that Perspective API can be implemented with acceptable levels of accuracy. However, the methodology revealed some misclassifications, highlighting the limitations of the Perspective API and the need for caution when using it in new research, something researchers should be mindful of when suggesting new methodological contributions (Bergh et al., 2022). Some scholars support these findings, arguing that the Perspective API is too general and fails to identify offensive terms targeted at specific minority groups, making it less reliable compared to hand-coded approaches (Kumar et al., 2021: 299; Czymara et al., 2022). The Perspective API struggles to comprehend language in specific contexts, leading to inaccurate scores. Even a minor change, like adding a full stop or the word "not", can alter the context and affect the scores (Hosseini et al., 2017). False alarms often occurred due to double negatives, as observed in Chapter Five, where some tweets were mistakenly classified as anti-refugee due to undetected double negatives by the classifier. This finding aligns with previous criticisms of the Perspective API classifier, which has shown vulnerability to false alarms and a tendency to identify words associated with identity attacks rather than specifically targeting offensive terms aimed at specific minority groups (Hosseini et al., 2017). Therefore, while the Perspective API classifier generally performs well, it still possesses significant limitations, such as the inability to detect double negatives, which should be carefully considered before making empirical claims.

Chapter Six presented novel insights into how NGOs can effectively navigate social media to promote cohesion in both digital and offline spaces. While this chapter provided valuable information, it had certain limitations. One limitation is that it did not explore how other stakeholders involved in cohesion utilise social media on a daily basis. Even though stakeholders like the police may not have dedicated social media accounts specifically for cohesion delivery, studying their use of social media in this context would still be informative. For instance, future research could investigate the frequency of cohesion-related tweets from these organisations in relation to their other core objectives, which would indicate the level of priority given to cohesion in their digital operations. Additionally, this chapter solely relied on three on-platform metrics (likes, retweets, and replies) to measure digital engagement (Boyd et al., 2010). Including other metrics, such as link clicks, to assess the effectiveness of pull strategy approaches or considering total donation figures

resulting from posts would provide a more comprehensive understanding of the differences in effectiveness between the identified strategies. Presenting a broader range of metrics would help avoid a skewed perception that dialogical and broadcasting approaches, such as push and networking strategies, are the most effective simply because they align more closely with engagement metrics like retweets and replies. Lastly, incorporating a language formality classifier would be beneficial for this chapter. Such a classifier could help determine whether professional and formal tones or informal and interpersonal tones are more effective in garnering public engagement. Although some anecdotal evidence in this chapter suggests that informal tones may be more effective in eliciting public engagement and cooperation, this relationship could be further explored in future studies by implementing a classifier or using manual coding methods.

## Bibliography

---

- Abburi, H., Sehgal, S., Maheshwari, H., & Varma, V. (2021). Knowledge-based Neural Framework for Sexism Detection and Classification. In *IberLEF@SEPLN* (pp. 402-414).
- Abdullah, A., R. Emery, C., & P. Jordan, L. (2020). Neighbourhood collective efficacy and protective effects on child maltreatment: A systematic literature review. *Health & Social Care in the Community*, 28(6), 1863-1883.
- Agarwal, D., Gelfand, A. and Citron-Pousty, S. (2002) 'Zero-inflated models with application to spatial count data', *Environmental and Ecological Statistics*, 9(4), pp. 341-355.
- Ahmed, W. (2018) 'Using social media data for research: An overview of tools', *Journal of Communication Technology*, 1(1), pp.77-94.
- Albornoz, F., Bradley, J., & Sonderegger, S. (2020). *The Brexit referendum and the rise in hate: Conforming to the new norm* (No. 2020-12). CeDEX Discussion Paper Series.
- Albors, J., Ramos, J. C., & Hervas, J. L. (2008). New learning network paradigms: Communities of objectives, crowd-sourcing, wikis and open source. *International Journal of Information Management*, 28, 194–202.
- Alfano, M., Reimann, R., Quintana, I. O., Chan, A., Cheong, M., & Klein, C. (2022). The Affiliative Use of Emoji and Hashtags in the Black Lives Matter Movement in Twitter. *Social Science Computer Review*, 08944393221131928.
- Allen, R., & Slotterback, C. S. (2017). Building immigrant engagement practice in urban planning: The case of Somali refugees in the Twin Cities. *Journal of Urban Affairs*, 43(6), 740-755.
- Allport, G. W. (1954). *The nature of prejudice*. Addison-Wesley.
- Almerekhi, H., Kwak, H., Salminen, J., & Jansen, B. J. (2020). Are these comments triggering? predicting triggers of toxicity in online discussions. In *Proceedings of the web conference 2020* (pp. 3033-3040).
- Alorainy, W., Burnap, P., Liu, H., & Williams, M. L. (2019). "The enemy among us" detecting cyber hate speech with threats-based othering language embeddings. *ACM Transactions on the Web (TWEB)*, 13(3), 1-26.

- Alves, H., Fernandes, C. and Raposo, M. (2016). "Social Media Marketing: A Literature Review and Implications." *Psychology & Marketing*, vol. 33, no. 9, pp. 1029-1038.
- Andersson, J.J. and Malm, A., (2006). 'Public-private partnerships and the challenge of critical infrastructure protection'. *International CIIP handbook*, 2, pp.139-166.
- Anger, I. and Kittl, C. (2011) 'Measuring influence on Twitter', *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies*, pp.1-4.
- Ansari, S. (2013). Social capital and collective efficacy: Resource and operating tools of community social control. *Journal of Theoretical & Philosophical Criminology*, 5(2).
- Antoine, J. Y., Villaneau, J., & Lefeuvre, A. (2014, April). Weighted Krippendorff's alpha is a more reliable metrics for multi-coders ordinal annotations: experimental studies on emotion, opinion and coreference annotation. In *EACL 2014* (pp. 10-p).
- Arnstein, S. (1969.) A ladder of citizen participation. *Journal of the American Planning Association*, 35(4), pp. 216–224.
- Atchison, C. (2000). Emerging styles of social control on the Internet: Justice denied. *Critical Criminology*, 9, 85-100.
- Awan, I., & Zempi, I. (2020). 'You all look the same': Non-Muslim men who suffer Islamophobic hate in the post-Brexit era. *European Journal of Criminology*, 17(5), 585-602.
- Ayanso, A., & Moyers, D. (2015). Social media use in the public sector. In *Digital Solutions for Contemporary Democracy and Government* (pp. 63-85). IGI Global.
- Aydin, E., & Ozeren, E. (2020). Inclusion and exclusion of sexual minorities at organisations: Evidence from LGBT NGOs in Turkey and the UK. *Journal of Organisational Change Management*, 33(3), 567-578.
- Backor, K., Golde, S., & Nie, N. (2007). Estimating survey fatigue in time use study. In *international association for time use research conference*. Washington, DC.
- Bagdikian, B. H. (2019). The new media monopoly.

- Bai, Q., Dan, Q., Mu, Z., & Yang, M. (2019). A systematic review of emoji: Current research and future
- Bakalis, C. (2018). Rethinking cyberhate laws. *Information & Communications Technology Law*, 27(1), 86-110.
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, 84(2), 191.
- Barbieri, F., & Camacho-Collados, J. (2018). How gender and skin tone modifiers affect emoji semantics in twitter. *The Association for Computational Linguistics*.
- Barrie, C. (2020). Searching racism after George Floyd. *Socius*, 6, 2378023120971507.
- Bartunek, J. M., & Rynes, S. L. (2010). The construction and contributions of "implications for practice": What's in them and what might they offer?. *Academy of management Learning & education*, 9(1), 100-117.
- Basit, T. (2003). Manual or electronic? The role of coding in qualitative data analysis. *Educational research*, 45(2), 143-154.
- Bastiaensens, S., Vandebosch, H., Poels, K., Van Cleemput, K., DeSmet, A., & De Bourdeaudhuij, I. (2014). Cyberbullying on social network sites. An experimental study into bystanders' behavioural intentions to help the victim or reinforce the bully. *Computers in Human Behavior*, 31, 259-271.
- Bauman, Z. (2013). *Liquid modernity*. John Wiley & Sons.
- Bayley, C., & French, S. (2008). Designing a participatory process for stakeholder involvement in a societal decision. *Group Decision and Negotiation*, 17(3), 195-210.
- Bayley, D. and Shearing, C. (1996) *The Future of Policing*. *Law and Society Review* 30(3) pp. 585-606
- Bayley, D. H. (2001). The new structure of policing: description, conceptualisation and research agenda.
- Bell, D. (1973). The coming of Post-Industrial Society. In *The Educational Forum* (Vol. 40, No. 4, pp. 574-579). Taylor & Francis.
- Benier, K. (2019). The people or the place? An analysis of the protective factors of hate crime in multiethnic neighbourhoods in Australia. *Australian Journal of Social Issues*, 54(2), 157-172.



- Bergh, D. D., Boyd, B. K., Byron, K., Gove, S., & Ketchen Jr, D. J. (2022). What constitutes a methodological contribution?. *Journal of Management*, 48(7), 1835-1848.
- Berman, S.J., Abraham, S., Battino, B., Shipnuck, L. and Neus, A. (2007) 'New business models for the new media world', *Strategy & Leadership*, 35(4), pp.23-30.
- Bernatzky, C., Costello, M., & Hawdon, J. (2022). Who produces online hate?: An examination of the effects of self-control, social structure, & social learning. *American journal of criminal justice*, 1-20.
- Bertot, J. C., Jaeger, P. T., & Grimes, J. M. (2012). 'Promoting transparency and accountability through ICTs, social media, and collaborative e-government', *Transforming government: people, process and policy*, vol. 6, no. 1, pp. 78-91.
- Bertot, John C., Jaeger, Paul T., & Grimes, Justin M. (2010). "Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption." *Government Information Quarterly*, vol. 27, pp. 263-271.
- Betts, R.K., (2002). *The trouble with strategy: Bridging policy and operations*. National Defense Univ Washington Dc Inst for National Strategic Studies.
- Bigg, M. (2008). Election of Obama provokes rise in US hate crimes. *Reuters*, November, 24.
- Billis, D. (2001). Tackling social exclusion: The contribution of voluntary organisations. In *Voluntary Organisations and social policy in Britain* (pp. 37-48). Palgrave, London.
- Billis, D., & Glennerster, H. (1998). Human services and the voluntary sector: towards a theory of comparative advantage. *Journal of Social policy*, 27(1), 79-98.
- Birnbacher, D. (1999). Ethics and social science: Which kind of co-operation?. *Ethical theory and moral practice*, 2, 319-336.
- Black, D. (1983). Crime as social control. *American sociological review*, 34-45.
- Blake, G., Diamond, J., Foot, J., Gidley, B., Mayo, M., Shukra, K., & Yarnit, M. (2008). *Community engagement and community cohesion*. Joseph Rowntree Foundation, York.
- Blalock, H. M. (1956). Economic discrimination and Negro increase. *American Sociological Review*, 21(5), 584-588.

- Bohning, D., Dietz, E., Schlattmann, P., Mendonca, L. and Kirchner, U. (1999) 'The zero-inflated Poisson model and the decayed, missing and filled teeth index in dental epidemiology', *Journal of the Royal Statistical Society. Series A*, pp. 195-209.
- Borgatti, S. P., & Everett, M. G. (2000). Models of core/periphery structures. *Social networks*, 21(4), 375-395.
- Bortree, D. S., & Seltzer, T. (2009). Dialogic strategies and outcomes: An analysis of environmental advocacy groups' Facebook profiles. *Public relations review*, 35(3), 317-319.
- Bowen, P., Rose, R., & Pilkington, A. (2017). Mixed methods-theory and practice. Sequential, explanatory approach. *International Journal of Quantitative and Qualitative Research Methods*, 5(2), 10-27.
- Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1), 210-230.
- Boyd, D., Golder, S., & Lotan, G. (2010). Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. In 2010 43rd Hawaii International Conference on System Sciences, 1-10. IEEE.
- Braga, A., Papachristos, A., & Hureau, D. (2012). Hot Spots Policing Effects on Crime. *Campbell Systematic Reviews*.
- Brainard, L. A., & McNutt, J. G. (2010). Virtual government–citizen relations: Informational, transactional, or collaborative?. *Administration & society*, 42(7), 836-858.
- Brantingham, P. J., Mohler, G., & MacDonald, J. (2022). Changes in public–police cooperation following the murder of George Floyd. *PNAS Nexus*, 1(5), pgac189.
- Bråten, S. (1981). Modeller av menneske og samfunn: Bro mellom teori og erfaring fra sosiologi og. *Oslo: Universitetsforlaget (in Norwegian)*.
- Braun, V. and Clarke, V. (2013), *Successful Qualitative Research*, 1st ed., SAGE Publications, London.
- Braun, V., Clarke, V., Boulton, E., Davey, L., & McEvoy, C. (2021). The online survey as a qualitative research tool. *International journal of social research methodology*, 24(6), 641-654.
- Bretschneider, S. I., & Mergel, I. (2010). Technology and public management information systems: Where have we been and where are we going. In D. C.

- Menzel, & H. J. White (Eds.), *The state of public administration: Issues, problems and challenges* (pp. 187–203). New York: M.E. Sharpe, Inc.
- Brett, M. (2022). Enabling cyber incident collaboration in UK local government through fast-time communication. *Cyber Security: A Peer-Reviewed Journal*, 5(3), 237-250.
- Briones, R. L., Kuch, B., Liu, B. F., & Jin, Y. (2011). Keeping up with the digital age: How the American Red Cross uses social media to build relationships. *Public relations review*, 37(1), 37-43.
- Brodie, E., Cowling, E., Nissen, N., Paine, A. E., Jochum, V., & Warburton, D. (2009). Understanding participation: A literature review.
- Brown, A. (2018). What is so special about online (as compared to offline) hate speech?. *Ethnicities*, 18(3), 297-326.
- Bruinsma, G. J., Pauwels, L. J., Weerman, F. M., & Bernasco, W. (2013). Social disorganisation, social capital, collective efficacy and the spatial distribution of crime and offenders: An empirical test of six neighbourhood models for a Dutch city. *British Journal of Criminology*, 53(5), 942-963.
- Bruns, A., & Burgess, J. (2011). The use of Twitter hashtags in the formation of ad hoc publics. In *Proceedings of the 6th European consortium for political research (ECPR) general conference 2011* (pp. 1-9). The European Consortium for Political Research (ECPR).
- Bryman, A., Becker, S., & Sempik, J. (2008). Quality criteria for quantitative, qualitative and mixed methods research: A view from social policy. *International journal of social research methodology*, 11(4), 261-276.
- Bullock, K. (2016), '(Re)presenting 'order' online: the construction of police presentational strategies on social media', *Policing and Society*, pp. 1-15.
- Bunar, N. (2007). Hate crimes against immigrants in Sweden and community responses. *American Behavioral Scientist*, 51(2), 166-181.
- Burch, L. F. (2020). *Experiences of everyday hate and hate crime in the lives of disabled people: meaning-making, impressions, and resistance* (Doctoral dissertation, University of Leeds).
- Burke, M., Kraut, R., & Marlow, C. (2011) Social capital on Facebook: Differentiating uses and users. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 571-580).

- Burnap, P., & Williams, M. L. (2015). Cyber hate speech on twitter: An application of machine classification and statistical modeling for policy and decision making. *Policy & internet*, 7(2), 223-242.
- Burnap, P., Rana, O. F., Avis, N., Williams, M., Housley, W., Edwards, A., ... & Sloan, L. (2015). Detecting tension in online communities with computational Twitter analysis. *Technological Forecasting and Social Change*, 95, 96-108.
- Burnap, P., Rana, O., Williams, M., Housley, W., Edwards, A., Morgan, J., Sloan, L., et al. (2014). 'COSMOS: Towards an integrated and scalable service for analysing social media on demand', *International Journal of Parallel, Emergent and Distributed Systems*, vol. 30, no. 2, pp. 80-100.
- Burnap, P., Rana, O., Williams, M., Housley, W., Edwards, A., Morgan, J., Sloan, L., et al. (2014), 'COSMOS: Towards an integrated and scalable service for analysing social media on demand', *International Journal of Parallel, Emergent and Distributed Systems*, Vol. 30 No. 2, pp. 80–100.
- Burnett, J. (2017). Racial violence and the Brexit state. *Race & class*, 58(4), 85-97.
- Burnett, R. (2015). Images and videos - media for communication. London: MENA Report.
- Burris, S. (2004). 'Nodal Governance'. *Australian Journal of Legal Philosophy* , pp. 2.
- Bursik Jr, R. J. Grasmick, HG (1993). Neighborhoods and crime: The dimensions of effective community control.
- Burston, P., Nfa, P. B., & Richardson, C. (Eds.). (2005). *A queer romance: Lesbians, gay men and popular culture*. Routledge.
- Bursztyń, L., Egorov, G., & Fiorin, S. (2017). *From extreme to mainstream: How social norms unravel* (No. w23415). National Bureau of Economic Research.
- Burton, S. and Soboleva, A. (2011) 'Interactive or reactive? Marketing with Twitter', *Journal of Consumer Marketing*, 28(7), pp.491-499.
- Burton-Smith, I., Sturgis, P., & Leckie, G. (2017). How collective is collective efficacy? The importance of consensus in judgments about community cohesion.
- Byers, B. D., & Jones, J. A. (2007). The impact of the terrorist attacks of 9/11 on anti-Islamic hate crime. *Journal of Ethnicity in Criminal Justice*, 5(1), 43-56.
- Cantle, T. (2018). *Community cohesion: A new framework for race and diversity*. Springer.

- Cao, X., Guo, X., Liu, H., & Gu, J. (2015). The role of social media in supporting knowledge integration: A social capital analysis. *Information Systems Frontiers*, 17, 351-362.
- Caracelli, V. J., & Greene, J. C. (1997). Crafting mixed-method evaluation designs. *New directions for evaluation*, 74, 19-32.
- Carletta, J. (1996). Assessing agreement on classification tasks: the kappa statistic. *arXiv preprint cmp-lg/9602004*.
- Carr, P. J. (2005). *Clean Streets: Controlling Crime, Maintaining Order, and Building Community Activism*. New York University Press. New York.
- Carron, A. V. (1982). Cohesiveness in sport groups: Interpretations and considerations. *Journal of Sport psychology*, 4(2).
- Casado-Molina, A. M., Rojas-de Gracia, M. M., Alarcón-Urbistondo, P., & Romero-Charneco, M. (2019). Exploring the opportunities of the emojis in brand communication: The case of the beer industry. *International Journal of Business Communication*, 59(3), 315-333.
- Castells, M. (1996), *The Information Age*, Blackwell, Oxford.
- Castells, M. (1997). An introduction to the information age. *City*, 2(7), 6-16.
- Castells, M. (2000a). Materials for an exploratory theory of the network society. *British Journal of Sociology*, 51(1), 5-24. Retrieved from <http://sociology.soc.uoc.gr/socmedia/zambarloukou/A001K/castells.pdf> 2.
- Castells, M. (2000b). *The rise of the network society* (2nd ed.). Blackwell Publishing. 3.
- Castells, M. (2009), *Communication Power*, Oxford University Press, Oxford.
- Castells, M. (2011). A Network Theory of Power. *International Journal of Communication*, 5, 773-787. Retrieved from <https://faculty.georgetown.edu/irvinem/theory/CastellsNetwork-Power-2011.pdf> 4.
- Castillo, M., Petrie, R., & Wardell, C. (2014). Fundraising through online social networks: A field experiment on peer-to-peer solicitation. *Journal of public economics*, 114, 29-35.
- Castronovo, C., & Huang, L. (2012). Social media in an alternative marketing communication model. *Journal of marketing development and competitiveness*, 6(1), 117-134.

- Ceruzzi, P.E. (2003), *A History of Modern Computing*, MIT Press, London, Eng.
- Chadwick Martin Bailey. (2011a). 10 quick facts you should know about consumer behavior on Facebook. [Online]. Available at: <http://forms.cmbinfo.com/10-quick-factsyou-should-know-about-consumer-behavior-on-facebook>
- Chadwick Martin Bailey. (2011b). 10 quick facts you should know about consumer behavior on Twitter. Available at [http://www.cmbinfo.com/cmb-cms/wp-content/uploads/2011/10/Twitter\\_Report\\_V11\\_2011.pdf](http://www.cmbinfo.com/cmb-cms/wp-content/uploads/2011/10/Twitter_Report_V11_2011.pdf)
- Chaffee, E. E. (1985). Three models of strategy. *Academy of management review*, 10(1), 89-98.
- Chakraborti, N. (2016). Mind the gap! Making stronger connections between hate crime policy and scholarship. *Criminal justice policy review*, 27(6), 577-589.
- Chakraborti, N. (2018). Responding to hate crime: Escalating problems, continued failings. *Criminology & Criminal Justice*, 18(4), 387-404.
- Chakraborti, N., & Garland, J. (2003). An 'invisible' problem? Uncovering the nature of racist victimisation in rural Suffolk. *International review of victimology*, 10(1), 1-17.
- Chan, J. and Bennett Moses, L. (2015), 'Is Big Data challenging criminology?', *Theoretical*
- Charities Against Hate. (2020). *A Guide to Best Practice in Ethical Digital Marketing & Comms Practices*. Retrieved from <https://img1.wsimg.com/blobby/go/09b24699-e035-4294-bb9b-49a9a58d43a7/Guide%20to%20Best%20Practice%20in%20Ethical%20Digital-0001.pdf>.
- Charities Against Hate. (2023). About us. [Online] Available at: <https://charitiesagainsthate.com/about-us> [Accessed: 25 January 2023].
- Chelsea, G., & Hansen, K. (2021). Did Covid-19 lead to an increase in hate crime against Chinese people in London?. *Journal of Contemporary Criminal Justice*.
- Cheminais, R. (2009) *Developing and evaluating multi-agency partnerships*. London: Routledge, pp.23-42.
- Chen, F. (2019). Leveraging interactive social media communication for organisational success: An examination of Chinese net-roots NGOs' microblog use. *Corporate Communications: An International Journal*, 24(1), 143-161.

- Chen, H. T. (2006). A theory-driven evaluation perspective on mixed methods research. *Research in the Schools*, 13(1), 75-83.
- Chetty, N., & Alathur, S. (2018). Hate speech review in the context of online social networks. *Aggression and violent behavior*, 40, 108-118. Chicago: University of Chicago Press.
- Cho, Y. (2008). *Encyclopaedia of Survey Research Methods*. Vol. 0. Thousand Oaks, CA: Sage Publications, Inc. Available at: <https://dx.doi.org/10.4135/9781412963947> [Accessed 6 Feb 2023].
- Choi, C. and Choi, S. (2012) 'Collaborative Partnerships and Crime in Disorganised Communities'. *Public Administration Review*, 72(2) pp.228-238.
- Choon, S. A., & Luna Reyes, L. F. (2012). "Social media in government." *Government Information Quarterly*, vol. 29, no. 4, pp. 441-445.
- Christensen, T., & Lægreid, P. (2006). The whole-of-government approach—regulation, performance, and public-sector reform.
- Chun, S. A., & Reyes, L. F. L. (2012). Social media in government. *Government Information Quarterly*, 29(4), 441-445.
- Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9), e2023301118. Clarendon Press, Oxford.
- Clarke, V., & Smith, M. (2015). "Not hiding, not shouting, just me": Gay men negotiate their visual identities. *Journal of Homosexuality*, 62(1), 4-32.
- Clear, T. R. (2009). *Imprisoning communities: How mass incarceration makes disadvantaged neighborhoods worse*. Oxford University Press.
- Cohen, A. K. (1955). *Delinquent boys; The culture of the gang*.
- Cohen, S. (1972). *Folk devils and moral panics: The creation of the mods and rockers. London, England*.
- Cohen, S. (2002). *Folk devils and moral panics: The creation of the mods and rockers. London, England*.
- Collins, C. R., Neal, Z. P., & Neal, J. W. (2017). Transforming social cohesion into informal social control: Deconstructing collective efficacy and the moderating role of neighborhood racial homogeneity. *Journal of Urban Affairs*, 39(3), 307-322.

- Corbin, J. and Strauss, A.L. (2015) *Basics of Qualitative Research*. Los Angeles: SAGE.
- Costello, M., Hawdon, J., & Ratliff, T. N. (2017). Confronting online extremism: The effect of self-help, collective efficacy, and guardianship on being a target for hate speech. *Social Science Computer Review*, 35(5), 587-605.
- Cotik, V., Debandi, N., Luque, F. M., Miguel, P., Moro, A., Pérez, J. M., ... & Zayat, D. (2020). A study of Hate Speech in Social Media during the COVID-19 outbreak.
- Craig, K. M. (2002). Examining hate-motivated aggression: A review of the social psychological literature on hate crimes as a distinct form of aggression. *Aggression and Violent Behavior*, 7(1), 85-101.
- Crandall, C. S., & Eshleman, A. (2003). A justification-suppression model of the expression and experience of prejudice. *Psychological bulletin*, 129(3), 414.
- Crandall, C. S., & Eshleman, A. (2005). The justification-suppression model of prejudice: An approach to the history of prejudice research. In *Social psychology of prejudice: historical and contemporary issues/edited by Christian S. Crandall, Mark Schaller*. Lawrence, Kan.: Lewinian Press, c2005.. Lawrence, Kan.: Lewinian Press.
- Crawford, A. (1997), *The Local Governance of Crime: Appeals to Community and Partnerships*,
- Crawford, A. (2002), "The governance of crime and insecurity in an anxious age: the trans-European and the local", in Crawford, A. (Ed.), *Crime and Insecurity*, Willan, Cullompton, pp. 27-51.
- Creswell, J. W., & Plano Clark, V. L. (2011). Choosing a mixed methods design. *Designing and conducting mixed methods research*, 2, 53-106.
- Creswell, J.W. (1994). *Research Design: Qualitative & Quantitative Approaches*. London: SAGE Publications.
- Creswell, J.W. and Plano Clark, V.L. (2018), *Designing and Conducting Mixed Methods Research*, SAGE Publications.
- Crime and Disorder Act (1998). HMSO, London. *Journal of Scandinavian Studies in Criminology and Crime Prevention*, 12(2), 198-215.
- Crooks, A., Croitoru, A., Stefanidis, A., & Radzikowski, J. (2013). # Earthquake: Twitter as a distributed sensor system. *Transactions in GIS*, 17(1), 124-147.



- Crump, J. (2011) 'What Are the Police Doing on Twitter? Social Media, the Police and the Public', *Policy & Internet*, 3(4), pp. 1-27.
- Cullen, A. L. (2022). Police and social media in England and Wales: Exploring police communications and image work in the digital society (Doctoral dissertation, Cardiff University, School of Social Sciences).
- Curtice, J. (2016). Brexit: Behind the referendum. *Political Insight*, 7(2), 4-7.
- Curtis, L., Edwards, C., Fraser, K. L., Gudelsky, S., Holmquist, J., Thornton, K., & Sweetser, K. D. (2010). Adoption of social media for public relations by nonprofit organisations. *Public relations review*, 36(1), 90-92.
- Curwen, P. (2022). Must I tweet? If you musk. *Digital Policy, Regulation and Governance*, 24(4), 398-399.
- Czymara, C. S., Dochow-Sondershaus, S., Drouhot, L. G., Simsek, M., & Spörlein, C. (2022). Catalyst of hate? Ethnic insulting on YouTube in the aftermath of terror attacks in France, Germany and the United Kingdom 2014–2017. *Journal of Ethnic and Migration Studies*, 1-19.
- Dadashzadeh, M. (2010). Social media in government. From eGovernment to eGovernance.
- Daniels, J. (2008). Race, civil rights, and hate speech in the digital era.
- Dann, S. (2015). Twitter data acquisition and analysis: Methodology and best practice. In *Maximizing Commerce and Marketing Strategies through Micro-Blogging* (pp. 280-296). IGI Global.
- Davenport, S.W., Bergman, S.M., Bergman, J.Z. and Fearington, M.E., 2014. Twitter versus Facebook: Exploring the role of narcissism in the motives and usage of different social media platforms. *Computers in human behavior*, 32, pp.212-220.
- De Widt, D., & Panagiotopoulos, P. (2018). Informal networking in the public sector: Mapping local government debates in a period of austerity. *Government Information Quarterly*, 35(3), 375-388.
- Defeis, E. F. (1992). Freedom of speech and international norms: A response to hate speech. *Stan. J. Int'l L.*, 29, 57.
- Del Rio-Chillcce, A., Jara-Monge, L., & Andrade-Arenas, L. (2021). Analysis of the use of videoconferencing in the learning process during the pandemic at a university in Lima. *International Journal of Advanced Computer Science and Applications*, 12(5).

- DeLeon-Granados, W. (1999). *Travels through crime and place: Community building as crime control*. UPNE.
- Demireva, N., & McNeil, R. (2015). Immigration, diversity and social cohesion. *Migration Observatory Briefing*.
- Denef, S., Bayerl, P.S. and Kaptein, N. (2013), 'Proceedings of the SIGCHI Conference on Human Factors in Computing Systems', in *Social media and the Police - Tweeting Practices of British Police Forces during the August 2011 Riots, Paris*, pp. 3471–3480.
- Deschamps, R., & McNutt, K. (2014). Third sector and social media. *Canadian journal of nonprofit and social economy research*, 5(2).
- Devine, D. (2018). The UK referendum on membership of the European Union as a trigger event for hate crimes. *Available at SSRN 3118190*.
- Dhakal, K. (2022). NVivo. *Journal of the Medical Library Association*, 110(2), 270-272.
- Dhaoui, C., Webster, C. M., & Tan, L. P. (2017). Social media sentiment analysis: lexicon versus machine learning. *Journal of Consumer Marketing*, 34(6), 480-488.
- Dillon, K. P., & Bushman, B. J. (2015). Unresponsive or un-noticed?: Cyberbystander intervention in an experimental cyberbullying context. *Computers in Human Behavior*, 45, 144-150.
- Dipoppa, G., Grossman, G., & Zonszein, S. (2022). Locked Down, Lashing Out: COVID-19 Effects on Asian Hate Crimes in Italy. *The Journal of Politics*, 85(2), 389-404.
- Dixon, S. (2018), 'Most popular social media platforms in the United Kingdom (UK) as of the third quarter 2021, by usage reach', Statista, available at: <https://www.statista.com/statistics/284506/united-kingdom-social-network-penetration/> (accessed 2022).
- Dorofeev, S. and Grant, P. (2006), *Statistics for Real-Life Sample Surveys: Non-simple Random Samples and Weighted Data*, Cambridge University Press, Cambridge.
- Douglas, K.M., McGarty, C., Bliuc, A.-M. & Lala, G. (2005). Understanding cyberhate: Social competition and social creativity in online white supremacist groups. *Social Science Computer Review*, 23(1), 68-76.

- Dovidio, J. F., Love, A., Schellhaas, F. M., & Hewstone, M. (2017). Reducing intergroup bias through intergroup contact: Twenty years of progress and future directions. *Group Processes & Intergroup Relations*, 20(5), 606-620.
- Doyle, L., Brady, A. M., & Byrne, G. (2009). An overview of mixed methods research. *Journal of research in nursing*, 14(2), 175-185.
- Drakulich, K. M., Crutchfield, R. D., Matsueda, R. L., & Rose, K. (2012). Instability, informal control, and criminogenic situations: Community effects of returning prisoners. *Crime, law and social change*, 57(5), 493-519.
- Drouhot, L. G., Deutschmann, E., Zuccotti, C. V., & Zagheni, E. (2022). Computational approaches to migration and integration research: promises and challenges. *Journal of Ethnic and Migration Studies*, 1-19.
- Dunn, B. (2010). "Best Buy's CEO on learning to love social media." *Harvard Business Review*, December, 43-48.
- Dupont, B. (2004), "Security in the age of networks", *Policing & Society*, Vol. 14 No. 1, pp. 76-91.
- Dutton, W. H. (1996). Network rules of order: Regulating speech in public electronic fora. *Media, Culture & Society*, 18(2), 269-290.
- Dutton, W. H., Fulk, J., & Steinfield, C. (1982). Utilisation of video conferencing. *Telecommunications Policy*, 6(3), 164-178.
- Dyfed Powys Police. (2008) Community Tension Force Policy Document.
- Earl, J., & Schussman, A. (2002). The new site of activism: On-line organisations, movement entrepreneurs, and the changing location of social movement decision making. In *Consensus decision making, Northern Ireland and indigenous movements* (pp. 155-187). Emerald Group Publishing Limited.
- Earl, J., McKee Hurwitz, H., Mejia Mesinas, A., Tolan, M., & Arlotti, A. (2013). This protest will be tweeted: Twitter and protest policing during the Pittsburgh G20. *Information, communication & society*, 16(4), 459-478.
- Edwards, A., Housley, W., Williams, M., Sloan, L., and Williams, M. (2013) 'Digital social research, social media and the sociological imagination: Surrogacy, augmentation and re-orientation', *International Journal of Social Research Methodology*, 16(3), pp. 245-260.
- Edwards, G. S., & Rushin, S. (2018). The effect of President Trump's election on hate crimes. *Available at SSRN 3102652*.

- Erickson, L. B. (2011). Social media, social capital, and seniors: The impact of Facebook on bonding and bridging social capital of individuals over 65.
- ESRC. (2019), "Research Ethics", *UKRI*, Economic and Social Research Council, available at: <https://esrc.ukri.org/funding/guidance-for-applicants/research-ethics/> (accessed 2019).
- European Commission (2017). European Digital Economy and Society Index (DESI). Retrieved from <https://ec.europa.eu/digital-single-market/en/desi> 5.
- Evans, J. R., & Mathur, A. (2005). The value of online surveys. *Internet research*, 15(2), 195-219.
- Evans, J. R., & Mathur, A. (2018). The value of online surveys: A look back and a look ahead. *Internet research*.
- Fan, L., Yu, H., & Gilliland, A. J. (2023). 1. AGGRAVATED ANTI-ASIAN HATE SINCE COVID-19 AND THE# STOPASIANHATE MOVEMENT: CONNECTION, DISJOINTNESS, AND CHALLENGES. *Hate Speech on Social Media*, 17.
- Farrimond, H. (2012). *Doing ethical research*. Bloomsbury Publishing.
- Feilzer, M. (2010). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of mixed methods research*, 4(1), 6-16.
- Ferguson, L. and Soave, V. (2020), '#Missing to #Found: Exploring police Twitter use for missing persons investigations', *Police Practice and Research*, pp. 1–17.
- Ferlander, S., & Timms, D. (1999). Social cohesion and online community.
- Festinger, L. (1950). Informal social communication. *Psychological review*, 57(5), 271.
- Festinger, L. (1952). Some consequences of deindividuation in a group. *Journal of Social Psychology*, 47, 382-389.
- Fieldhouse, E., & Cutts, D. (2008). Diversity, density and turnout: The effect of neighbourhood ethno-religious composition on voter turnout in Britain. *Political Geography*, 27(5), 530-548.
- Fielding, N. G. (1995). *Community policing*. Oxford University Press.

- Fielding, N., & Innes, M. (2006). Reassurance policing, community policing and measuring police performance. *Policing & society*, 16(02), 127-145.
- Fiske, S. T. (2002). What we know now about bias and intergroup conflict, the problem of the century. *Current Directions in Psychological Science*, 11(4), 123-128.
- Ford, R. A. (2004). *Local diversity, prejudice and white majority welfare attitudes in the UK* (Doctoral dissertation, Nuffield College). formal social control approach, police institutions also need to use an informal social control approach by mobilising community social agents (Pramono, 2020:102).
- Forrest, R., & Kearns, A. (2001). Social cohesion, social capital and the neighbourhood. *Urban studies*, 38(12), 2125-2143.
- Friedmann, J. (1967). A general theory of polarised development.
- Froehlich, D. E., Rehm, M., & Rienties, B. C. (2020). Mixed methods social network analysis. *Mixed methods social network analysis: Theories and methodologies in learning and education*, 1-10.
- Fuchs, C. (2012). Social media, riots, and revolutions. *Capital & Class*, 36(3), 383-391.
- Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). *Countering online hate speech*. UNESCO Publishing.
- Gallacher, J. D. (2021). Leveraging cross-platform data to improve automated hate speech detection. *arXiv preprint arXiv:2102.04895*.
- Gallagher, R. J., Young, J. G., & Welles, B. F. (2021). A clarified typology of core-periphery structure in networks. *Science advances*, 7(12), eabc9800.
- Ganis, M., & Kohirkar, A. (2015). *Social media analytics: Techniques and insights for extracting business value out of social media*. IBM Press.
- Garay, A. M., Hashimoto, E. M., Ortega, E. M., & Lachos, V. H. (2011). On estimation and influence diagnostics for zero-inflated negative binomial regression models. *Computational Statistics & Data Analysis*, 55(3), 1304-1318.
- Garland, D. (1996). THE LIMITS OF THE SOVEREIGN STATE Strategies of Crime Control in Contemporary Society. *The British journal of criminology*, 36(4), 445-471.
- Garland, D. (2001), *The Culture of Control*, Oxford University Press, Oxford

- Garland, J., & Chakraborti, N. (2004). Racist Victimisation, Community Safety and the Rural: Issues and Challenges. *British Journal of Community Justice*, 2, 21-32.
- Garland, J., & Chakraborti, N. (2006). 'Race', space and place: Examining identity and cultures of exclusion in rural England. *Ethnicities*, 6(2), 159-177.
- Garland, J., Ghazi-Zahedi, K., Young, J. G., Hébert-Dufresne, L., & Galesic, M. (2022). Impact and dynamics of hate and counter-speech online. *EPJ data science*, 11(1), 3.
- Garnham, N. (1998). Information society theory as ideology: A critique. *Loisir et société/Society and Leisure*, 21(1), 97-120.
- Gau, J. M. (2014). Unpacking collective efficacy: The relationship between social cohesion and informal social control. *Criminal Justice Studies*, 27(2), 210-225.
- Gau, J. M., & Brunson, R. K. (2010). Procedural justice and order maintenance policing: A study of inner-city young men's perceptions of police legitimacy. *Justice quarterly*, 27(2), 255-279.
- Gearhart, M. C., & Joseph, M. L. (2019). Social cohesion, mutual efficacy, and informal social control: Enhancing the conceptualisation of collective efficacy. *Community Development*, 50(1), 3-15.
- Gerstenfeld, P.B. (2017). *Hate crimes: Causes, controls, and controversies*. Sage.
- Giani, M., & Méon, P. G. (2021). Global racist contagion following Donald Trump's election. *British Journal of Political Science*, 51(3), 1332-1339.
- Giddens A. (1990), *The Consequences of Modernity*. Stanford: Stanford University Press
- Gil de Zúñiga, H., Jung, N., & Valenzuela, S. (2012). Social media use for news and individuals' social capital, civic engagement and political participation. *Journal of computer-mediated communication*, 17(3), 319-336.
- Gil-García, J. R. (2006). Enacting state websites: A mixed method study exploring e-government success in multi-organisational settings. In *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)* (Vol. 4, pp. 76b-76b). IEEE.
- Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, 457, 1012–1014.

- Glaeser, E. L. (2005). The political economy of hatred. *The Quarterly Journal of Economics*, 120(1), 45-86.
- Glassman, M., Kuznetcova, I., Peri, J., & Kim, Y. (2021). Cohesion, collaboration and the struggle of creating online learning communities: Development and validation of an online collective efficacy scale. *Computers and Education Open*, 2, 100031.
- Gligorić, K., Anderson, A., and West, R. (2018) 'How constraints affect content: The case of Twitter's switch from 140 to 280 characters', In *Proceedings of the International AAI Conference on Web and Social Media*, vol. 12, no. 1.
- Goldsmith, A. (2015). Disgracebook policing: social media and the rise of police indiscretion. *Policing and society*, 25(3), 249-267.
- Goldsmith, A.J. (2010). "Policing's New Visibility." *British Journal of Criminology*, vol. 50, no. 5, pp. 914-934.
- Goodhart, D. (2004). "Too diverse?" *Prospect*, 95, 30-37.
- Google. (2021a) 'About the API', Perspective API. Available at: <https://support.perspectiveapi.com/s/about-the-api> (Accessed: 13 February 2022).
- Google. (2021b) 'About the Scores', Perspective API. Available at: [https://developers.perspectiveapi.com/s/about-the-api-score?language=en\\_US](https://developers.perspectiveapi.com/s/about-the-api-score?language=en_US) (Accessed: 13 February 2022).
- Gover, A. R., Harper, S. B., & Langton, L. (2020). Anti-Asian hate crime during the COVID-19 pandemic: Exploring the reproduction of inequality. *American journal of criminal justice*, 45, 647-667.
- Granovetter, M. S. (1973). The strength of weak ties. *American journal of sociology*, 78(6), 1360-1380.
- Grant, S. (2009). Silos of helplessness: A cautionary tale about introducing social media (Web 2.0) applications in the community sector. *Third sector Review*, 15, 89-104.
- Gray, L. M., Wong-Wylie, G., Rempel, G. R., & Cook, K. (2020). Expanding qualitative research interviewing strategies: Zoom video communications. *The Qualitative Report*, 25(5), 1292-1301.
- Green, D. P., McFalls, L. H., & Smith, J. K. (2001). Hate crime: An emergent research agenda. *Annual review of sociology*, 27(1), 479-504.

- Greenberg, J., & MacAulay, M. (2009). NPO 2.0? Exploring the web presence of environmental nonprofit organisations in Canada. *Global Media Journal: Canadian Edition*, 2(1).
- Groff, E. R. (2015). Informal social control and crime events. *Journal of Contemporary Criminal Justice*, 31(1), 90-106.
- Gronholm, P. C., Henderson, C., Deb, T., & Thornicroft, G. (2017). Interventions to reduce discrimination and stigma: the state of the art. *Social psychiatry and psychiatric epidemiology*, 52(3), 249-258.
- Gruver, M. (2019). Reforming Police in The Post-George Floyd Era: The Role of Public Opinion and Bureaucratic Police Community in Reform Outcomes.
- Gruzd, A., & Mai, P. (2021). Commanalytic: A Research Tool for Studying Online Communities and Online Discourse. Available at <https://Commanalytic.com/> (Accessed: 13 February 2022).
- Guest, G. (2013). Describing mixed methods research: An alternative to typologies. *Journal of mixed methods research*, 7(2), 141-151.
- Guimarães, S. S., Reis, J. C., Ribeiro, F. N., & Benevenuto, F. (2020). Characterising toxicity on facebook comments in brazil. In *Proceedings of the Brazilian Symposium on Multimedia and the Web* (pp. 253-260).
- Guo, C., & Saxton, G. D. (2010). Voice-in, voice-out: Constituent participation and nonprofit advocacy. In *Nonprofit policy forum* (Vol. 1, No. 1). De Gruyter.
- Guo, C., & Saxton, G. D. (2014). Tweeting social change: How social media are changing nonprofit advocacy. *Nonprofit and voluntary sector quarterly*, 43(1), 57-79.
- Guo, C., & Saxton, G. D. (2018). Speaking and being heard: How nonprofit advocacy organisations gain attention on social media. *Nonprofit and Voluntary Sector Quarterly*, 47(1), 5-26.
- Gwet, K. L. (2011). On the Krippendorff's alpha coefficient. *Manuscript submitted for publication*. Retrieved October, 2(2011), 2011.
- Hall, L. L. (2018). *Race and Online Hate: Exploring the Relationship between Race and the Likelihood of Exposure to Hate Material Online* (Doctoral dissertation, Virginia Tech).
- Halpern, D. (2005). *Social capital*. Polity.



- Halpin, D. R., Fraussen, B., & Ackland, R. (2021). Which audiences engage with advocacy groups on Twitter? Explaining the online engagement of elite, peer, and mass audiences with advocacy groups. *Nonprofit and Voluntary Sector Quarterly*, 50(4), 842-865.
- Han, X., Gu, X., and Peng, S. (2019) Analysis of Tweet Form's effect on users' engagement on Twitter, *Cogent Business & Management*.
- Hanes, E. and Machin, S. (2014) 'Hate Crime in the Wake of Terror Attacks: Evidence from 7/7 and 9/11', *Journal of Contemporary Criminal Justice*, 30, pp. 247-267.
- Hankes, K., & Zhang, S. (2017). A Waning Storm: Once the World's Most Popular White Nationalist Website, Stormfront is Running out of Steam.
- Hanna, R., Rohm, A., & Crittenden, V. L. (2011). We're all connected: The power of the social media ecosystem. *Business horizons*, 54(3), 265-273.
- Hanzelka, J., & Schmidt, I. (2017). Dynamics of Cyber Hate in Social Media: A Comparative Analysis of Anti-Muslim Movements in the Czech Republic and Germany. *International Journal of Cyber Criminology*, 11(1).
- Hardy, S. J., & Chakraborti, N. (2020). *The Process of Engagement: Hard to Reach or Easy to Ignore?* Palgrave Hate Studies. Palgrave Pivot, Cham.
- Harel, T. O., Jameson, J. K., & Maoz, I. (2020). The normalisation of hatred: Identity, affective polarisation, and dehumanisation on Facebook in the context of intractable political conflict. *Social Media+ Society*, 6(2), 2056305120913983.
- Hart, T. C., & Colavito, V. (2011). College student victims and reporting crime to the police: The influence of collective efficacy. *Criminology, Criminal Justice, Law & Society*, 12(3), 1.
- Hassan, G., Rabah, J., Madriaza, P., Brouillette-Alarie, S., Borokhovski, E., Pickup, D., ... & Danis, E. (2022). PROTOCOL: Hate online and in traditional media: A systematic review of the evidence for associations or impacts on individuals, audiences, and communities. *Campbell Systematic Reviews*, 18(2), e1245.
- Hässler, T., Uluğ, Ö. M., Kappmeier, M., & Travaglino, G. A. (2021). Intergroup contact and social change: An integrated contact-collective action model. *Journal of Social Issues*, 77(1), 217-241.
- Haukoos, J. S., & Lewis, R. J. (2005). Advanced statistics: bootstrapping confidence intervals for statistics with "difficult" distributions. *Academic emergency medicine*, 12(4), 360-365.

- Hawdon, J., & Ryan, J. (2009). Social capital, social control, and changes in victimization rates. *Crime & delinquency*, 55(4), 526-549.
- Hawdon, J., Costello, M., Barrett-Fox, R., & Bernatzky, C. (2018). The Perpetuation of Online Hate: A Criminological Analysis of Factors Associated with Participating in a Online Attack. *J. Hate Stud.*, 15, 157.
- Hawdon, J., Oksanen, A., & Räsänen, P. (2015). Online extremism and online hate. *Nordicom Inf*, 37, 29-37.
- Hayes, J. L., Britt, B. C., Applequist, J., Ramirez Jr, A., & Hill, J. (2020). Leveraging textual paralinguage and consumer–brand relationships for more relatable online brand communication: a social presence approach. *Journal of Interactive Advertising*, 20(1), 17-30.
- Haythornthwaite C (2005) Social networks and Internet connectivity effects. *Information, Communication & Society* 8(2): 125–147.
- Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert-scales?: The reference-group effect. *Journal of personality and social psychology*, 82(6), 903.
- Hermes, J. & Hill, A. (2021). “Transgression in contemporary media culture”. *International Journal of Cultural Studies*, 24(1): 3-14. <https://doi.org/10.1177/1367877920968105>
- Hemsley, J. and Mason, R.M. (2013), ‘Knowledge and Knowledge Management in the Social Media Age’, *Journal of Organisational Computing and Electronic Commerce*, Vol. 23 No. 1-2, pp. 138–167.
- Hendricks, J. A., & Schill, D. (2017). The social media election of 2016. *The 2016 US presidential campaign: Political communication and practice*, 121-150.
- Hewstone, M. (2015). Consequences of diversity for social cohesion and prejudice: The missing dimension of intergroup contact. *Journal of Social Issues*, 71(2), 417-438.
- Higgins, B. R., & Hunt, J. (2016). Collective efficacy: Taking action to improve neighborhoods. *NIJ Journal*, 277, 18-21.
- Hipp, J. R. (2016). Collective efficacy: How is it conceptualised, how is it measured, and does it really matter for understanding perceived neighborhood crime and disorder?. *Journal of criminal justice*, 46, 32-44.
- Hipp, J. R., & Wo, J. C. (2015). Collective efficacy and crime. *International encyclopedia of social and behavioral sciences*, 4, 169-173.

- Hirschi, T. (1969). *Causes of delinquency*. Berkeley: University of California Press.
- Hirvonen, K. (2013). Sweden: When hate becomes the norm. *Race & Class*, 55(1), 78-86.
- Hoagwood, K. E., Atkins, M., Horwitz, S., Kutash, K., Olin, S. S., Burns, B., ... & Kelleher, K. J. (2018). A response to proposed budget cuts affecting children's mental health: Protecting policies and programs that promote collective efficacy. *Psychiatric services*, 69(3), 268-273.
- Hobolt, S. B. (2016). The Brexit vote: a divided nation, a divided continent. *Journal of European public policy*, 23(9), 1259-1277.
- Hodge R, Kress G (1988) *Social Semiotics*. Cambridge: Polity.
- Hodge R, Tripp D (1986) *Children and Television: A Semiotic Approach*. Cambridge: Polity.
- Hodgkinson, T. (2022). Crime and victimisation in rural areas. *Rural Victims of Crime: Representations, Realities and Responses*, 223.
- Hoffman, A. J. (2020). Community service activities reducing hate crimes and extremism: A "green intervention" approach. *Journal of Prevention & Intervention in the Community*, 48(3), 207-209.
- Holding, H., & Martin, C. (1998). Excel Functions. In *Mastering Microsoft® Office* (pp. 231-239). Palgrave, London.
- Hood, C. (1991). A public management for all seasons?. *Public administration*, 69(1), 3-19.
- Hooghe, M., Reeskens, T., Stolle, D., & Trappers, A. (2009). Ethnic diversity and generalised trust in Europe: A cross-national multilevel study. *Comparative political studies*, 42(2), 198-223.
- Hopkins, D. J. (2010). Politicized places: Explaining where and when immigrants provoke local opposition. *American political science review*, 104(1), 40-60.
- Hopkins, D. J., & Washington, S. (2020). The rise of Trump, the fall of prejudice? Tracking white Americans' racial attitudes via a panel survey, 2008–2018. *Public Opinion Quarterly*, 84(1), 119-140.
- Housley, W., Edwards, A., Williams, M., & Williams, M. (2013). Special Issue, Computational Social Science: Research, Design and Methods. *International Journal of Social Research Methods*, 16(3).

- Howard, P. N., Duffy, A., Freelon, D., Hussain, M. M., Mari, W., & Maziad, M. (2011). Opening closed regimes: what was the role of social media during the Arab Spring?. Available at SSRN 2595096.
- Hrdinová, J., Helbig, N., & Peters, C. S. (2010). Designing social media policy for government: Eight essential elements. Albany, NY: Center for Technology in Government, University at Albany.
- Hughes, J. A., & Sharrock, W. W. (2016). *The philosophy of social research*. Routledge. Human Behavior 32(1): 212–220.
- Hunt, J. S. (2007). Implicit bias and hate crimes: A psychological framework and Critical Race Theory analysis. In *Social Consciousness in Legal Decision Making* (pp. 247-265). Springer, Boston, MA.
- Hwang, H., & Kim, K. O. (2015). Social media as a tool for social movements: The effect of social media use and social capital on intention to participate in social movements. *International Journal of Consumer Studies*, 39(5), 478-488.
- Ireland, L., Hawdon, J., Huang, B., & Peguero, A. (2020). Preconditions for guardianship interventions in cyberbullying: Incident interpretation, collective and automated efficacy, and relative popularity of bullies. *Computers in human behavior*, 113, 106506.
- Iwama, J. A. (2018). Understanding hate crimes against immigrants: Considerations for future research. *Sociology compass*, 12(3), e12565.
- Jaishankar, K. (2007). Establishing a theory of cyber crimes. *International Journal of Cyber Criminology*, 1(2), 7-9.
- Janowitz, M. (1975). Sociological theory and social control. *American Journal of sociology*, 81(1), 82-108.
- John, A. S. (2020). It's not just Zoom. Google Meet, Microsoft Teams, and Webex have privacy issues too. *Consumer Reports*, 3.
- Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of mixed methods research*, 1(2), 112-133.
- Johnston, L. and Shearing, C. (2003), *The Governance of Security*, Routledge, London.
- Joinson, A. (1998). Causes and implications of disinhibited behaviour on the internet. In Gackenbach, J. (Ed.), *Psychology of the Internet: Intrapersonal, interpersonal, and transpersonal implications*.

- Jue, A.L., Marr, J.A. and Kassotakis, M.E. (2010), *Social Media at Work*, Jossey-Bass, San
- Kagan, C., & Duggan, K. (2011). Creating community cohesion: the power of using innovative methods to facilitate engagement and genuine partnership. *Social Policy and Society*, 10(3), 393-404.
- Kaplan, A. M., & Haenlein, M. (2009). The fairyland of Second Life: Virtual social worlds and how to use them. *Business horizons*, 52(6), 563-572.
- Kaplan, A.M., & Haenlein, M. (2010). Users of the world, unite! The challenges and
- Karapanos E, Teixeira P and Gouveia R (2016) Need fulfillment and experiences on social media: a case on Facebook and WhatsApp. *Computers in Human Behavior* 55(1): 888–897.
- Kavanaugh, A., Fox, E. A., Sheetz, S., Yang, S., Li, L. T., Whalen, T., ... & Xie, L. (2012). Social media use by government: From the routine to the critical. In *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times* (pp. 121-130).
- Kean, T. H., & Hamilton, L. H. (2004). *The 9/11 Report*. New York: St. Martin's Press.
- Kelly, J. (2007). Reforming public services in the UK: Bringing in the third sector. *Public administration*, 85(4), 1003-1022.
- Kelly, L. M., & Cordeiro, M. (2020). Three principles of pragmatism for research on organisational processes. *Methodological innovations*, 13(2), 2059799120937242.
- Kendall, J., & Knapp, M. (1997). 'The United Kingdom', in L.M. Salamon, & H.K. Anheier (Eds.), *Defining the Nonprofit Sector*. Manchester: Manchester University Press.
- Kickert, W. J., Klijn, E. H., & Koppenjan, J. F. M. (Eds.). (1997). *Managing complex networks: Strategies for the public sector*. Sage.
- Kietzmann, J. H., Hermkens, K., McCarthy, I. P., and Silvestre, B. S. (2011) 'Social Media? Get Serious! Understanding the Functional Building Blocks of Social Media', *Business Horizons*, 54 (3), pp.241–51.
- Kim, B. J. (2015). Political efficacy, community collective efficacy, trust and extroversion in the information society: Differences between online and offline civic/political activities. *Government Information Quarterly*, 32(1), 43-51.

- King, R.D. and Sutton, G.M. (2013) 'High times for hate crimes: Explaining the temporal clustering of hate-motivated offending', *Criminology*, 51(4), pp. 871-894.
- Kleinberg, S. (2013). To hashtag or not to hashtag?: That is the question-how to use them correctly without going overboard. *Chicago Tribune; Chicago, Ill*, 2-5.
- Knott, K., & Poole, E. (2016). *Media portrayals of religion and the secular sacred: Representation and change*. Routledge.
- Kochel, T. R. (2012). Can police legitimacy promote collective efficacy?. *Justice quarterly*, 29(3), 384-419.
- Koopmans, R., & Olzak, S. (2004). Discursive opportunities and the evolution of right-wing violence in Germany. *American journal of Sociology*, 110(1), 198-230.
- Kopytowska, M., & Baider, F. (2017). From stereotypes and prejudice to verbal and physical violence: Hate speech in context. *Lodz Papers in Pragmatics*, 13(2), 133-152.
- Kornhauser, R. R. (1978). Social sources of delinquency: An appraisal of analytic models.
- Kruskal, J.B. and Wish, M. (1978), 'Multidimensional Scaling', Sage University Paper Series on Quantitative Applications in the Social Sciences, Sage, London.
- Kumar, D., Kelley, P. G., Consolvo, S., Mason, J., Bursztein, E., Durumeric, Z., ... & Bailey, M. (2021). Designing Toxic Content Classification for a Diversity of Perspectives. In *SOUPS@USENIX Security Symposium* (pp. 299-318).
- Kurbin, C., Weitzer, R. (2003). "New Directions in Social Disorganisation Theory. *George Washington University*.
- Kwon, E. S., & Sung, Y. (2011). Follow me! Global marketers' Twitter use. *Journal of Interactive Advertising*, 12(1), 4-16.
- Kyrollos, D. G., & Green, J. R. (2021). MetaHate: A meta-model for hate speech detection. In *2021 IEEE International Conference on Big Data (Big Data)* (pp. 2496-2502). IEEE.
- Lai, Z., Liu, W., Wu, Q., Li, H., Xu, J., & Wu, J. (2022). SPACERTC: Unleashing the low-latency potential of mega-constellations for real-time communications. In *IEEE INFOCOM 2022-IEEE Conference on Computer Communications*, 1339-1348. IEEE.

- Lapinski MK and Rimal RN (2005) An explication of social norms. *Communication Theory* 15(2): 127–147.
- Lash, S. (2001) 'Technological Forms of Life', *Theory, Culture and Society* 18(1): 105–20
- Laurence, J., & Heath, A. (2008). Predictors of community cohesion: multi-level modelling of the 2005 Citizenship Survey.
- Lee, M. and McGovern, A., (2013). *Policing and media: Public relations, simulations and communications*. Routledge.
- Legewie, J. (2013). Terrorist Events and Attitudes Toward Immigrants: A Natural Experiment. *American Journal of Sociology*, 118, 1199-1245.
- Legg, S. (2021). 'No place for hate': community-led research and the geographies of Nottingham citizens' hate crime commission. *Social & Cultural Geography*, 22(8), 1164-1186.
- Lehman, B. (2019). Stopping the hate: Applying insights on bullying victimisation to understand and reduce the emergence of hate in schools. *Sociological inquiry*, 89(3), 532-555.
- Leishman, F. and Mason, P. (2003), *Policing and the Media*, 1st ed., Routledge Taylor & Francis
- Lepineux, F. (2005). Stakeholder theory, society and social cohesion. *Corporate Governance: The international journal of business in society*.
- Letki, N. (2008). Does diversity erode social cohesion? Social capital and race in British neighbourhoods. *Political Studies*, 56(1), 99-126.
- Levi, M. and Leighton Williams, M. (2013) 'Multi-agency partnerships in cybercrime reduction'. *Information Management & Computer Security*, 21(5) pp.420-443.
- Levin J., Rabrenovic G. (2004). *Why we hate*. Amherst, NY: Prometheus.
- Levin, J., & McDevitt, J. (2002). *Hate crimes revisited: America's war on those who are different*. Westview Press.
- Levine, J. M., & Moreland, R. L. (1994). Group socialization: Theory and research. *European review of social psychology*, 5(1), 305-336.

- Lickel, B., Miller, N., Stenstrom, D. M., Denson, T. F., & Schmader, T. (2006). Vicarious retribution: The role of collective blame in intergroup aggression. *Personality and social psychology review*, 10(4), 372-390.
- Lieberman, J.D., Koetzle, D. and Sakiyama, M. (2013). 'Police Departments' Use of Facebook', *Police Quarterly*, 16(4), pp. 438-462.
- Lima, L., Reis, J. C., Melo, P., Murai, F., & Benevenuto, F. (2020, December). Characterising (un) moderated textual data in social systems. In *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (pp. 430-434). IEEE.
- Lin, H. C., Swarna, H., & Bruning, P. F. (2017). Taking a global view on brand post popularity: Six social media brand post practices for global markets. *Business Horizons*, 60(5), 621-633.
- Lin, H. C., Swarna, H., & Bruning, P. F. (2017). Taking a global view on brand post popularity: Six social media brand post practices for global markets. *Business Horizons*, 60(5), 621-633.
- Lindblad, M. R., Manturuk, K. R., & Quercia, R. G. (2013). Sense of community and informal social control among lower income households: The role of homeownership and collective efficacy in reducing subjective neighborhood crime and disorder. *American journal of community psychology*, 51(1), 123-139.
- Lindgren, S. (2017), *Digital Media and Society*, 1st ed., Sage Publications, London.
- Livingston, M., Bailey, N., & Kearns, A. (2008). People's attachment to place: The influence of neighbourhood deprivation.
- Lombard, M., Snyder-Duch, J. and Bracken, C.C. (2002), 'Content Analysis in Mass Communication: Assessment and Reporting of Intercoder Reliability', *Human Communication Research*, Vol. 28 No. 4, pp. 587–604.
- Lomborg, S., & Bechmann, A. (2014). Using APIs for data collection on social media. *The Information Society*, 30(4), 256-265.
- Lovejoy, K., & Saxton, G. D. (2012). Information, community, and action: How nonprofit organisations use social media. *Journal of computer-mediated communication*, 17(3), 337-353.
- Luhmann, N. (1996). *Die Realität der Massenmedien*. Opladen, Germany: Westdeutscher Verlag.



- Lukka, V., & James, P. T. (2014). Attitudes toward Facebook advertising. *Journal of management and Marketing Research*, 14, 1.
- Lup K, Trub L and Rosenthal L (2015) Instagram #Instasad?: Exploring associations among Instagram use, depressive symptoms, negative social comparison, and strangers followed. *Cyberpsychology, Behavior, and Social Networking* 18(5): 247–252.
- Lynch, R. (2001). An analysis of the concept of community cohesion. *Community Cohesion: A Report of the Independent Review Team*, 69-76.
- Ma, L. (2013). Diffusion and assimilation of government microblogging. *Public Management Review*, <http://dx.doi.org/10.1080/14719037.2012.725763>.
- Macaulay, S. (1986). Private Government'in Law and the Social Sciences, eds. L. Lipson and S. Wheeler, 445.
- Macdonald, S. J., Donovan, C., & Clayton, J. (2017). The disability bias: understanding the context of hate in comparison with other minority populations. *Disability & Society*, 32(4), 483-499.
- Malcom, Z. T., Wenger, M. R., & Lantz, B. (2022). Politics or prejudice? Separating the influence of political affiliation and prejudicial attitudes in determining support for hate crime law. *Psychology, Public Policy, and Law*.
- Mangold, W. G., & Faulds, D. J. (2009). "Social media: The new hybrid element of the promotion mix." *Business Horizons*, vol. 52, pp. 357-365.
- Manikonda, L., Hu, Y., & Kambhampati, S. (2014). Analysing user activities, demographics, social network structure and user-generated content on Instagram. arXiv preprint arXiv:1410.8099.
- Manjubashini, B., Murali, K., Sujith, C., Vignesh Kumar, J., & Vigneshwaran, P (2022). Detection of Cyberbullying on Social Media.
- Mansourifar, H., Alsagheer, D., Fathi, R., Shi, W., Ni, L., & Huang, Y. (2022). Hate speech detection in clubhouse. In *Machine Learning and Principles and Practice of Knowledge Discovery in Databases: International Workshops of ECML PKDD 2021, Virtual Event, September 13-17, 2021, Proceedings, Part II* (pp. 341-351). Cham: Springer International Publishing.
- Markudova, D., Trevisan, M., Garza, P., Meo, M., Munafò, M. M., and Carofiglio, G. (2021) 'What's my app? ml-based classification of rtc applications', *ACM SIGMETRICS Performance Evaluation Review*, 48(4), pp.41-44.
- Marres, N. (2018), *Digital Sociology*, Polity, Cambridge

- Marwick AE and Boyd D (2011) I tweet honestly, I tweet passionately: twitter users, context collapse, and the imagined audience. *New Media & Society* 13(1): 114–133.
- Mawby, R.C. (1998), 'Policing the image', *Criminal Justice Matters*, Vol. 32, pp. 26–7.
- Mawby, R.C. (2002), *Policing Images: Policing, Communication and Legitimacy*, 1st ed., Willan, Cullompton.
- Mazerolle, L., Wickes, R., & McBroom, J. (2010). Community variations in violence: The role of social ties and collective efficacy in comparative context. *Journal of Research in Crime and Delinquency*, 47(1), 3-30.
- McKinsey (2015a) *Digital America: A Tale of the Haves and Have-Mores*.
- McLaren, L. M. (2003). Anti-immigrant prejudice in Europe: Contact, threat perception, and preferences for the exclusion of migrants. *Social forces*, 81(3), 909-936.
- McLaughlin C and Vitak J (2012) Norm evolution and violation on Facebook. *New Media & Society* 14(2): 299–315.
- McLeod, J., & Von Treuer, K. (2013). Towards a cohesive theory of cohesion.
- Meijer, A., & Thaens, M. (2013). Social media strategies: Understanding the differences between North American police departments. *Government information quarterly*, 30(4), 343-350.
- Meijer, A., Koops, B. J., Pieterse, W., Overman, S., & Ten Tije, S. (2012). "Government 2.0: Key challenges to its realization." *Electronic Journal of e-Government*, 10, 59-69.
- Meraji Oskuie, S., Mohamadkhani, K., Delavar, A., & Farhangi, A. A. (2022). A Path Analysis of Factors Affecting Social Control of Cybercultural Transgressions. *Journal of Cyberspace Studies*, 6(1), 47-70.
- Mergel, I. (2012) 'The social media innovation challenge in the public sector', *Information Polity*, 17, pp. 281-292.
- Mergel, I. (2013). A framework for interpreting social media interactions in the public sector. *Government information quarterly*, 30(4), 327-334.
- Mergel, I., Mugar, G., & Hossein Jarrahi, M. (2012). Forming and norming social media adoption in the corporate sector Proceedings iConference 2012, February 7–10, Toronto, Canada.

- Merton, R. K. (1957). Social structure and anomie. *American sociological review*, 3(5), 672-682.
- Meyers, L. S., Gamst, G. C., & Guarino, A. J. (2013). *Performing data analysis using IBM SPSS*. John Wiley & Sons.
- Millie, A., & Herrington, V. (2005). Bridging the gap: Understanding reassurance policing. *The Howard Journal of Criminal Justice*, 44(1), 41-56.
- Mintzberg, H. (1978). Patterns in strategy formation. *Management science*, 24(9), 934-948.
- Mintzberg, H., & Waters, J. A. (1985). Of strategies, deliberate and emergent. *Strategic management journal*, 6(3), 257-272.
- Monnier, J., Knapp, R. G., & Frueh, B. C. (2003). Recent advances in telepsychiatry: an updated review. *Psychiatric Services*, 54(12), 1604-1609.
- Montague, I., Gazal, K. A., Wiedenbeck, J., & Shepherd, J.-G. (2016). "Forest products industry in a digital age: A look at e-commerce and social media." *Forest Products Journal*, 66, 49-57. doi:10.13073/FPJ-D-14-00104
- Moore, M., & Ramsay, G. (2017). *UK media coverage of the 2016 EU Referendum campaign*. London: King's College London.
- Moore, N. (1997). The information society. *World information report*, 98, 271-284.
- Moreland, R. L., & Levine, J. M. (1982). Socialization in small groups: Temporal changes in individual-group relations. In *Advances in experimental social psychology* (Vol. 15, pp. 137-192). Academic Press.
- Morenoff, J. D., Sampson, R. J., & Raudenbush, S. W. (2001). Neighborhood inequality, collective efficacy, and the spatial dynamics of urban violence. *Criminology*, 39(3), 517-558.
- Morgan, D.L. (2013), *Integrating Qualitative and Quantitative Methods*, SAGE Publications, Inc.
- Morse, J. M. (2016). *Mixed method design: Principles and procedures*. Routledge.
- Mossberger, K., Tolbert, C. J., & McNeal, R. S. (2008). Digital citizenship: The internet, society,
- Mossberger, K., Wu, Y., & Crawford, J. (2013). Connecting citizens and local governments? Social media and interactivity in major US cities. *Government information quarterly*, 30(4), 351-358.

- Mowlabocus, S. (2016). *Gaydar culture: Gay men, technology and embodiment in the digital age*. Routledge.
- Müller, K., & Schwarz, C. (2021). Fanning the flames of hate: Social media and hate crime. *Journal of the European Economic Association*, 19(4), 2131-2167.
- Murphy, A. (2022). *How does political rhetoric influence hate speech?* (Doctoral dissertation, University of Leicester).
- Myers, W., & Lantz, B. (2020). Reporting racist hate crime victimisation to the police in the United States and the United Kingdom: A cross-national comparison. *The British Journal of Criminology*, 60(4), 1034-1055.
- Myhill, A. (2006) *Community Engagement in Policing: Lessons from the literature*. London: Home Office.
- NatCen (2014) *Research Using Social Media: Users' Views*, London: Natcen.
- Naurin, D., & Lindahl, R. (2008). East-North-South: Coalition-building in the Council before and after enlargement. *Unveiling the Council of the European Union: Games governments play in Brussels*, 64-78.
- Naveed, N., Gottron, T., Kunegis, J., & Alhadi, A. C. (2011, June). Bad news travel fast: A content-based analysis of interestingness on twitter. In *Proceedings of the 3rd international web science conference* (pp. 1-7).
- Nayak, M. S. D. P., & Narayan, K. A. (2019). Strengths and weaknesses of online surveys. *Technology*, 6(7), 0837-2405053138.
- Niedzwiecka, M., & Pan, Y. Ch. (2017). An Exploratory Study into Employee Attitudes towards Digitalisation of Library Services in Higher Education. *Proceedings of the 22nd UK Academy for Information Systems Annual Conference (UKAIS 2017) April 2017, Oxford, UK*. 7.
- Nistico, A., Markudova, D., Trevisan, M., Meo, M., & Carofiglio, G. (2020). A comparative study of RTC applications. In *2020 IEEE International Symposium on Multimedia (ISM)*, 1-8. IEEE.
- Nix, J., Wolfe, S. E., Rojek, J., & Kaminski, R. J. (2015). Trust in the police: The influence of procedural justice and perceived collective efficacy. *Crime & Delinquency*, 61(4), 610-640.
- Nolan, J. J., & Akiyama, Y. (1999). An analysis of factors that affect law enforcement participation in hate crime reporting. *Journal of Contemporary Criminal Justice*, 15(1), 111-127.

- O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: debates and practical guidelines. *International journal of qualitative methods*, 19, 1609406919899220.
- O'Connor, C. D. (2015). The police on Twitter: image management, community building, and implications for policing in Canada. *Policing and society*, 27(8), 899-912.
- Obar, J. A., & Wildman, S. S. (2015). Social media definition and the governance challenge-an introduction to the special issue. Obar, JA and Wildman, S. (2015). Social media definition and the governance challenge: An introduction to the special issue. *Telecommunications policy*, 39(9), 745-750.
- Oberwittler, D. (2007). The effects of neighbourhood poverty on adolescent problem behaviours: A multi-level analysis differentiated by gender and ethnicity. *Housing Studies*, 22(5), 781-803.
- Oboler, A. (2016). Measuring the hate. In *The State of Antisemitism in Social Media*. Online Hate Prevention Institute. Produced for the Global Forum for Combating Antisemitism.
- Odgers, C. L., Moffitt, T. E., Tach, L. M., Sampson, R. J., Taylor, A., Matthews, C. L., & Caspi, A. (2009). The protective effects of neighborhood collective efficacy on British children growing up in deprivation: a developmental analysis. *Developmental psychology*, 45(4), 942.
- Okeowo, A. (2016). Hate on the rise after Trump's election. *The New Yorker*, 17.
- Oksanen, A., Kaakinen, M., Minkkinen, J., Räsänen, P., Enjolras, B., & Steen-Johnsen, K. (2020). Perceived societal fear and cyberhate after the November 2015 Paris terrorist attacks. *Terrorism and Political Violence*, 32(5), 1047-1066.
- Oldenburg, R. (1999) *The Great Good Place*. New York: Marlowe & Company.
- Oldendick, R. W. (2012). Survey research ethics. *Handbook of survey methodology for the social sciences*, 23-35.
- Olinski, M., & Szamrowski, P. (2021). Facebook as an engagement tool: How are public benefit organisations building relationships with their public?. *Plos one*, 16(9), e0256880.
- Oliver, E., & Wong, J. (2003). Intergroup prejudice in multiethnic settings. *American journal of political science*, 47(4), 567-582.
- Olivet, J., McGraw, S., Grandin, M., & Bassuk, E. (2010). Staffing challenges and strategies for organisations serving individuals who have experienced chronic

- homelessness. *The journal of behavioral health services & research*, 37, 226-238.
- Olzak, S. (1990). The political context of competition: Lynching and urban racial violence, 1882–1914. *Social forces*, 69(2), 395-421.
- ONS. (2020), 'Internet access – households and individuals, Great Britain', Ons.gov.uk, Office for National Statistics, available at: <https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/bulletins/internetaccesshouseholdsandindividuals/2020>.
- O'reilly, T., 2008. What Is Web 2.0-Design patterns and business models for the next generation of software. <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>.
- Orellana, C. I. (2023). Health workers as hate crimes targets during COVID-19 outbreak in the Americas. *Revista de salud publica*, 22, 253-257.
- Orton-Johnson. K and N. Prior (Ed) *Digital Sociology: Critical Perspectives*. New York: Palgrave.
- Ozalp, S., Williams, M. L., Burnap, P., Liu, H., & Mostafa, M. (2020). Antisemitism on Twitter: Collective efficacy and the role of community organisations in challenging online hate speech. *Social Media+ Society*, 6(2), 2056305120916850.
- Page, B., Ames, A., Pask, R., & Palenicek, A. (2004). *What Drives Public Satisfaction with Local Government*. Local Government Research Report 20/04), LGA Publications, London.
- Panagiotopoulos, P., Bigdeli, A.Z. and Sams, S. (2014), 'Citizen–government collaboration on social media: The case of Twitter in the 2011 riots in England', *Government Information Quarterly*, Vol. 31 No. 3, pp. 349–357.
- Pandey, N., and Bein, D. (2018) 'Web application for social networking using RTC', In *IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 336-340. IEEE.
- Park, H.W., Barnett, G.A. and Nam, I.Y. (2002) 'Interorganisational hyperlink networks among websites in South Korea', *NETCOM: Réseaux, communication et territoires/Networks and Communication Studies*, 16(3), pp.155-173.
- Parkes, J., Abercrombie, S., & McCarty, T. (2013). Feedback sandwiches affect perceptions but not performance. *Advances in Health Sciences Education*, 18, 397-407.

- Pass, M. D., Madon, N. S., Murphy, K., & Sargeant, E. (2020). To trust or distrust?: Unpacking ethnic minority immigrants' trust in police. *The British Journal of Criminology*, 60(5), 1320-1341.
- Pastor-Escuredo, D., & Vinuesa, R. (2020). Towards and Ethical Framework in the Complex Digital Era. *arXiv preprint arXiv:2010.10028*.
- Patton, M. Q. (2002). Two decades of developments in qualitative inquiry: A personal, experiential perspective. *Qualitative social work*, 1(3), 261-283.
- Pennant, R. (2005). *Diversity, trust and community participation in England*. Home Office. Research, Development and Statistics Directorate.
- Percy-Smith, J. (2003). In search of social capital. *Policy & Politics*, 31(3), 321-334.
- Perna, G., Markudova, D., Trevisan, M., Garza, P., Meo, M., Munafò, M. M., and Carofiglio, G. (2021) 'Online classification of RTC traffic', In IEEE 18th Annual Consumer Communications & Networking Conference (CCNC), pp. 1-6. IEEE.
- Perry, B. (2001). *In the name of hate: Understanding hate crimes*. Psychology Press.
- Perry, B., & Olsson, P. (2009). Cyberhate: the globalization of hate. *Information & Communications Technology Law*, 18(2), 185-199.
- Perry, B., Mirrlees, T., & Scrivens, R. (2017). The Dangers of Porous Borders: The "Trump Effect" in Canada. *J. Hate Stud.*, 14, 53.
- Perry, B., Perry, J., Schweppe, J., & Walters, M. (2016). Introduction: Understanding hate crime: Research, policy, and practice. *Criminal justice policy review*, 27(6), 571-576.
- Petrosino, C., & Pace, J. (2015). Social cohesion, collective efficacy, and the response of a Cape Verdean community to hate crime: Learning a new reality. *American Behavioral Scientist*, 59(13), 1681-1697.
- Piatkowska, S. J., & Lantz, B. (2021). Temporal clustering of hate crimes in the aftermath of the Brexit vote and terrorist attacks: A comparison of Scotland and England and Wales. *The British Journal of Criminology*, 61(3), 648-669.
- Piatkowska, S. J., & Stults, B. J. (2022). Brexit, terrorist attacks, and hate crime: A longitudinal analysis. *Social Problems*, 69(4), 968-996.
- Plan, E. (2014). Modeling and simulation of count data. *CPT: pharmacometrics & systems pharmacology*, 3(8), 1-12.

- Poister, T. H., & Thomas, J. C. (2007). "The wisdom of crowds: Learning from administrators' predictions of citizen perceptions." *Public Administration Review*, 67(2), 279-289.
- Porter M. E. (2001). "Strategy and the Internet." *Harvard Business Review*, March 2001.
- Porter, M. E. (1980). *Competitive Strategy: Techniques for Analysing Industries and Competitors*, Free Press, New York, NY.
- Posner, P. L. (2009). The pracademic: An agenda for re-engaging practitioners and academics. *Public Budgeting & Finance*, 29(1), 12-26.
- Potter, W. J., & Levine-Donnerstein, D. (1999). Rethinking validity and reliability in content analysis.
- Powell, A., Stratton, G., & Cameron, R. (2018). *Digital criminology: Crime and justice in digital society*. Routledge.
- Powell, S. G. (2000). Specialisation, teamwork, and production efficiency. *International Journal of Production Economics*, 67(3), 205-218.
- Pramono, G. E. (2020). A Formal Form of Social Control Against Online Hate Speech In Indonesia. *Social Sciences and Education Research Review*, 7(2), 89-106.
- Presdee, M. (2000) *Cultural Criminology and the Carnival of Crime*. London: Routledge.
- Procter, R. and Voss, A., (2013). Reading the riots on Twitter: methodological innovation for the analysis of big data. In: W. Housley, A. Edwards, M.L. Williams, and M. Williams, eds. *Computational social science: research, design and methods* [Special Issue]. *International Journal of Social Research Methods*, 16 (3)
- Purcell, K., Rainie, L., & Brenner, J. (2012). *Search engine use 2012*.
- Purdy, C. H. (2011). Using the Internet and social media to promote condom use in Turkey. *Reproductive health matters*, 19(37), 157-165.
- Putnam, R. D. (2007). E pluribus unum: Diversity and community in the twenty-first century the 2006 Johan Skytte Prize Lecture. *Scandinavian political studies*, 30(2), 137-174.



- Quan-Haase A and Young AL (2010) Uses and gratifications of social media: a comparison of Facebook and instant messaging. *Bulletin of Science, Technology & Society* 30(5): 350–361.
- R. Ellis, J. (2022). A fairy tale gone wrong: Social media, recursive hate and the politicisation of Drag Queen Storytime. *The journal of criminal law*, 86(2), 94-108.
- Ragins, B. R., Singh, R., & Cornwell, J. M. (2007). Making the invisible visible: fear and disclosure of
- Raskin, R. (2006). Facebook faces its future. *Young Consumers*.
- Ratcliffe, J. (2003). *Intelligence-led policing* (Vol. 248). Canberra: Australian Institute of Criminology.
- Ravelli, L. J., & Van Leeuwen, T. (2018). Modality in the digital age. *Visual communication*, 17(3), 277-297.
- Regmi, P. R., Waithaka, E., Paudyal, A., Simkhada, P., & van Teijlingen, E. (2016). Guide to the design and application of online questionnaire surveys. *Nepal journal of epidemiology*, 6(4), 640
- Reiss, A. J. (1951). Delinquency as the failure of personal and social controls. *American sociological review*, 16(2), 196-207.
- Reny, T. T., & Newman, B. J. (2021). The opinion-mobilising effect of social protest against police violence: Evidence from the 2020 George Floyd protests. *American Political Science Review*, 115(4), 1499-1507.
- Rheingold, H. (1993) *The Virtual Community Homesteading on the Electronic Frontier*. New York: Harper Collins.
- Robles, P. (2011) 'Twitter isn't very social: study', Econsultancy, [blog].
- Rose, D. R., & Clear, T. R. (1998). Incarceration, social capital, and crime: Implications for social disorganisation theory. *Criminology*, 36, 441-480.
- Roxell, L. (2011). Hate, threats, and violence. A register study of persons suspected of hate
- Sadique, K., Tangen, J., & Perowne, A. (2018). The importance of narrative in responding to hate incidents following 'trigger'events. *UK: Leicester*.

- Şahi, H., Kılıç, Y., & Sağlam, R. B. (2018). Automated detection of hate speech towards woman on Twitter. In *2018 3rd international conference on computer science and engineering (UBMK)* (pp. 533-536). IEEE.
- Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). Earthquake shakes twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World wide web* (pp. 851-860).
- Salamon, L. M., & Anheier, H. K. (1997). *Defining the nonprofit sector: A cross-national analysis*. Manchester University Press.
- Salehabadi, N., Groggel, A., Singhal, M., Roy, S. S., & Nilizadeh, S. (2022). User Engagement and the Toxicity of Tweets. *arXiv preprint arXiv:2211.03856*.
- Salminen, J., Hopf, M., Chowdhury, S. A., Jung, S. G., Almerakhi, H., & Jansen, B. J. (2020). Developing an online hate classifier for multiple social media platforms. *Human-centric Computing and Information Sciences, 10*, 1-34.
- Salmon, G. (2004) 'Multi-Agency Collaboration: The Challenges for CAMHS'. *Child and Adolescent Mental Health, 9*(4) pp.156-161.
- Sampson, J. (2010). Collective Efficacy Theory. In *Encyclopedia of Criminological Theory*. Volume 2, eds. Francis T. Cullen and Pamela Wilcox, Sage Publications.
- Sampson, R. J. (2001). Crime and public safety: Insights from community-level perspectives on social capital. *Social capital and poor communities, 3*.
- Sampson, R. J. (2002). Transcending tradition: New directions in community research, Chicago style. *Criminology, 40*(2), 213-230.
- Sampson, R. J. (2012). *Great American City: Chicago and the Enduring Neighborhood Effect*.
- Sampson, R. J. (2017). Collective efficacy theory: Lessons learned and directions for future inquiry. *Taking stock, 149-167*.
- Sampson, R. J., & Raudenbush, S. W. (1999). Systematic social observation of public spaces: A new look at disorder in urban neighborhoods. *American journal of sociology, 105*(3), 603-651.
- Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *science, 277*(5328), 918-924.

- Sampson, R.J., (2001). Crime and public safety: Insights from community-level perspectives on social capital. In Saegert, S., Thompson, J.P. & Warren, M.R. (Eds.), *Social capital and poor communities* (pp. 89-114). Russell Sage.
- Sandholtz, N., Langton, L., & Planty, M. (2013). *Hate crime victimisation, 2003-2011*. Washington, DC: US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.
- Santhanam, S., Srinivasan, V., Mahajan, K., & Shaikh, S. (2021). Towards Understanding How Emojis Express Solidarity in Crisis Events. In *Advances in Artificial Intelligence, Software and Systems Engineering: Proceedings of the AHFE 2021 Virtual Conferences on Human Factors in Software and Systems Engineering, Artificial Intelligence and Social Computing, and Energy, July 25-29, 2021, USA* (pp. 133-140). Springer International Publishing.
- Sargeant, E., Wickes, R., & Mazerolle, L. (2013). Policing community problems: Exploring the role of formal social control in shaping collective efficacy. *Australian & New Zealand journal of criminology*, 46(1), 70-87.
- Savić, D. (2020). COVID-19 and work from home: Digital transformation of the workforce. *Grey Journal (TGJ)*, 16(2), 101-104.
- Saxton, G. D., Niyirora, J., Guo, C., & Waters, R. (2015). # AdvocatingForChange: The strategic use of hashtags in social media advocacy. *Advances in Social Work*, 16(1), 154-169.
- Sayer, A. (1992). *Method in social science: A realist approach* (2nd ed.). London: Routledge.
- Scheider, M. C., Chapman, R., & Schapiro, A. (2009). Towards the unification of policing innovations under community policing. *Policing: An international journal of police strategies & management*.
- Scheuermann, L., & Taylor, G. (1997). Netiquette. *Internet Research*, 7(4), 269-273.
- Schneider, C.J. (2016), *Policing and Social Media*, 1st ed., Lexington Books, Lanham.
- Schoonenboom, J., & Johnson, R. B. (2017). How to construct a mixed methods research design. *Kolner Zeitschrift fur Soziologie und Sozialpsychologie*, 69(Suppl 2), 107.
- Scott, E. (2020). Fact file: Rural economy. In *Focus. Agriculture & Farming. Economy*. House of Lords Library, UK Parliament, Westminster. Published Monday, 27 January, 2020.

- Schumann, S., & Moore, Y. (2023). The COVID-19 outbreak as a trigger event for sinophobic hate crimes in the United Kingdom. *The British Journal of Criminology*, 63(2), 367-383.
- Schweppe, J., Haynes, A., & Carr, J. (2014). A life free from fear: Legislating for hate crime in Ireland: An NGO perspective. University of Limerick
- Schweppe, J., & Perry, B. (2022). A continuum of hate: delimiting the field of hate studies. *Crime, Law and Social Change*, 77(5), 503-528.
- Schweppe, J., Haynes, A., & MacIntosh, E. M. (2020). What is measured matters: The value of third party hate crime monitoring. *European Journal on Criminal Policy and Research*, 26, 39-59.
- Schweppe, J., & Walters, M. A. (Eds.). (2016). *The globalization of hate: internationalizing hate crime?*. Oxford University Press.
- Scrivens, R., Burruss, G. W., Holt, T. J., Chermak, S. M., Freilich, J. D., & Frank, R. (2021). Triggered by defeat or victory? Assessing the impact of presidential election results on extreme right-wing mobilisation online. *Deviant Behavior*, 42(5), 630-645.
- Sedgwick, D., Callahan, J. and Hawdon, J. (2020) 'Institutionalizing partnerships: a mixed methods approach to identifying trends and perceptions of community policing and multi-agency task forces'. *Police Practice and Research*, 22(1) pp.727-744.
- Seegerberg, A. and Bennett, W.L. (2011) 'Social media and the organisation of collective action: using Twitter to explore the ecologies of two climate change protests', *The Communication Review*, 14(3), pp.197-215.
- Sela, S., Kuflik, T., & Mesch, G. S. (2012). Changes in the discourse of online hate blogs: The effect of Barack Obama's election in 2008. *First Monday*.
- Sevilla, G. (2020). Zoom vs. Microsoft Teams vs. Google Meet: Which top videoconferencing app is best. *PC Mag*.sexual orientation at work. *Journal of applied psychology*, 92(4), 1103.
- Shahidullah, Shahid. (2018). The Rise of the Digital Society, Generation Z, and Management Challenges in the 21st Century: A Review Essay. *Annals of Social Sciences & Management studies*. 1. 10.19080/ASM.2018.01.555551.
- Shakespeare, T., Watson, N., Brunner, R., Cullingworth, J., Hameed, S., Scherer, N., ... & Reichenberger, V. (2022). Disabled people in Britain and the impact of the COVID-19 pandemic. *Social Policy & Administration*, 56(1), 103-117.

- Shea, V. (1994). Core Rules of Netiquette. *Educom Review*, 29(5), 58-62.
- Shearing, C. (2001) 'A nodal conception of governance: Thoughts on a policing commission'. *Policing and Society*, 11(3-4) pp.259-272.
- Shearing, C., & Wood, J. (2003). Nodal governance, democracy, and the new 'denizens'. *Journal of law and society*, 30(3), 400-419.
- Sherman, E. (2018). A genius way to get more attention for tweets also makes life easier. Inc
- Shi, Z., Rui, H., & Whinston, A. B. (2014). Content sharing in a social broadcasting environment: evidence from twitter. *MIS quarterly*, 38(1), 123-142.
- Shirky, C. (2011). The political power of social media. *Foreign Affairs*, 90, 28-41.
- Siegel, A. A., Nikitin, E., Barberá, P., Sterling, J., Pullen, B., Bonneau, R., ... & Tucker, J. A. (2021). Trumping hate on Twitter? Online hate speech in the 2016 US election campaign and its aftermath. *Quarterly Journal of Political Science*, 16(1), 71-104.
- Silver, E., & Miller, L. L. (2004). Sources of informal social control in Chica neighborhoods. *Criminology*, 42(3), 551-584.
- Simi, P. (2010). Why study white supremacist terror? A research note. *Deviant Behavior*, 31(3), 251-273.
- Simons, R. L., Simons, L. G., Burt, C. H., Brody, G. H., & Cutrona, C. (2005). Collective efficacy, authoritative parenting and delinquency: A longitudinal test of a model integrating community-and family-level processes. *Criminology*, 43(4), 989-1029.
- Sin, C. H. (2013). 16 Hate crime against people with disabilities. *The Routledge international handbook on hate crime*.
- Singh, R., & Awasthi, S. (2020). Updated comparative analysis on video conferencing platforms-zoom, google meet, microsoft teams, webex teams and gotomeetings. *EasyChair Preprint*, 4026, 1-9.
- Sinickas, A. (2007). Finding a cure for survey fatigue. *Strategic Communication Management*, 11(2), 11.
- Skitka, L. J., Bauman, C. W., & Mullen, E. (2004). Political tolerance and coming to psychological closure following the September 11, 2001, terrorist attacks: An integrative approach. *Personality and Social Psychology Bulletin*, 30(6), 743-756.

- Skogan, W. G. (2006). Asymmetry in the impact of encounters with police. *Policing & society*, 16(02), 99-126.
- Slevitch, L. (2011). Qualitative and quantitative methodologies compared: Ontological and epistemological perspectives. *Journal of quality assurance in hospitality & tourism*, 12(1), 73-81.
- Sloan, L., & Morgan, J. (2015). Who tweets with their location? Understanding the relationship between demographic characteristics and the use of geoservices and geotagging on Twitter. *PLoS one*, 10(11), e0142209.
- Sloan, L., Morgan, J., Housley, W., Williams, M., Edwards, A., Burnap, P., & Rana, O. (2013). Knowing the Tweeters: Deriving sociologically relevant demographics from Twitter. *Sociological research online*, 18(3), 74-84.
- Slocum, A. L. E. E., Taylor, T. J., Brick, B. T., & Esbensen, F. A. (2010). Neighbourhood structural characteristics, individual-level attitudes, and youths' crime reporting intentions. *Criminology*, 48(4), 1063-1100.
- Slotterback, C. S., & Lauria, M. (2019). Building a foundation for public engagement in planning: 50 years of impact, interpretation, and inspiration from Arnstein's Ladder. *Journal of the American Planning Association*, 85(3), 183-187.
- Snelson, C. L. (2016). Qualitative and mixed methods social media research: A review of the literature. *International Journal of Qualitative Methods*, 15(1), 1609406915624574.
- Soares, F. B., Recuero, R., & Zago, G. (2019). Asymmetric polarisation on Twitter and the 2018 Brazilian presidential elections. In *Proceedings of the 10th international conference on social media and society* (pp. 67-76).
- Social Data Science Lab. (2021). *Ethics*, available at: <http://socialdatalab.net/ethics> (accessed 2019).
- South, L., Saffo, D., Vitek, O., Dunne, C., & Borkin, M. A. (2022, June). Effective use of Likert-scales in visualisation evaluations: a systematic review. In *Computer Graphics Forum* (Vol. 41, No. 3, pp. 43-55).
- Spelman, W. (1987). *Problem-oriented policing*. US Department of Justice, National Institute of Justice.
- Statista. (2022), 'Active social media audience in the United Kingdom (UK) in February 2022', [Staista.com](https://www.statista.com/statistics/507405/uk-active-social-media-and-mobile-social-media-users/), available at: <https://www.statista.com/statistics/507405/uk-active-social-media-and-mobile-social-media-users/>

- Stelzner, M. A. (2015). 2015 social media marketing industry report. Available at: <https://www.socialmediaexaminer.com/SocialMediaMarketingIndustryReport2015.pdf> [Accessed: 12th December 2022].
- Stemler, S. (2000). An overview of content analysis. *Practical assessment, research, and evaluation*, 7(1), 17.
- Stephan, W. G., Renfro, C. L., & Davis, M. D. (2008). The role of threat in intergroup relations. *Improving intergroup relations: Building on the legacy of Thomas F. Pettigrew*, 55-72.
- Sternberg, R. J., & Sternberg, K. (2008). *The nature of hate*. Cambridge University Press.
- Stewart, H., Mason, R., &. (2016). Theresa May in 'U-turn' over pre-article 50 Brexit debate in parliament. *The Guardian*, 12.
- Stolle, D., Petermann, S., Schmid, K., Schönwälder, K., Hewstone, M., Vertovec, S., ... & Heywood, J. (2013). Immigration-related diversity and trust in German cities: The role of intergroup contact. *Journal of Elections, Public Opinion & Parties*, 23(3), 279-298.
- Strathern, W., & Pfeffer, J. (2022). Identifying Different Layers of Online Misogyny. *arXiv preprint arXiv:2212.00480*.
- Stratton G, Powell A and Cameron R (2017) Crime and Justice in Digital Society: Towards a 'Digital Criminology'? *International Journal for Crime, Justice and Social Democracy* 6(2): 17-33. DOI: 10.5204/ijcjsd.v6i2.355.
- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research*. Sage publications.
- Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A., Nadal, K. L., & Esquilin, M. (2007). Racial microaggressions in everyday life: implications for clinical practice. *American psychologist*, 62(4), 271.
- Suhr, D.D., (2005). Principal component analysis vs. exploratory factor analysis. *SUGI 30 proceedings*, 203, p.230.
- Sullaway, M. (2004). Psychological perspectives on hate crime laws. *Psychology, Public Policy, and Law*, 10(3), 250.
- Sunshine, J., & Tyler, T. R. (2003). The role of procedural justice and legitimacy in shaping public support for policing. *Law & society review*, 37(3), 513-548.
- Sunstein, C.R. (2017). *#Republic: Divided democracy in the age of social media*. Princeton University Press.

- Sutter, O. (2019). Narratives of “Welcome Culture”: the cultural politics of voluntary aid for refugees. *Narrative Culture*, 6(1), 19-43.
- Swartz, M., Crooks, A., & Kennedy, W. (2020). Diversity from emojis and keywords in social media. In *International Conference on Social Media and Society* (pp. 92-100).
- Symantec (2017a) Internet Security Threat Report, Vol (23).
- Syrdal, H. A., & Briggs, E. (2018). Engagement with social media content: A qualitative exploration. *Journal of Marketing Theory & Practice*, 26,4–22.
- Tabachnick, B.G. and Fidell, L.S. (2013), *Using Multivariate Statistics*, 6th ed., Allyn & Bacon, Boston, MA.
- Tajfel, H., Turner, J. C., Austin, W. G., & Worchel, S. (1979). An integrative theory of intergroup conflict. *Organisational identity: A reader*, 56(65), 9780203505984-16.
- Tashakkori, A., & Teddlie, C. (2003). The past and future of mixed methods research: From data triangulation to mixed model designs. *Handbook of mixed methods in social and behavioral research*, 671, 701.
- Tausch, N., Tam, T., Hewstone, M., Kenworthy, J., & Cairns, E. (2007). Individual-level and group-level mediators of contact effects in Northern Ireland: The moderating role of social identification. *British journal of social psychology*, 46(3), 541-556.
- Teddlie, C., & Tashakkori, A. (2011). Mixed methods research. *The Sage handbook of qualitative research*, 4, 285-300.
- Tell MAMA. (2017). *Beyond the incident: Outcomes for victims of anti-Muslim prejudice*. Technical Report. <https://tellmamauk.org/wp-content/uploads/2018/07/Tell-MAMA-Report-2017.pdf>.
- Tereseviciene, M., Tamoliune, G., Naujokaitiene, J., Pranckute, D., & Ehlers, U. D. (2018). Characteristics of digital and network society: emerging places and spaces of learning. In *EDEN Conference Proceedings* (No. 1, pp. 19-21).
- Tetlock, P. E., Kristel, O. V., Elson, S. B., Green, M. C., & Lerner, J. S. (2000). The psychology of the unthinkable: taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of personality and social psychology*, 78(5), 853.
- Thew, D. (2008). LinkedIn—a user's perspective: Using new channels for effective business networking. *Business information review*, 25(2), 87-90.



- Thomas, P. (2011). *Youth, multiculturalism and community cohesion*. Springer.
- Treadwell, J., & Garland, J. (2011). Masculinity, marginalisation and violence: A case study of the English Defence League. *The British Journal of Criminology*, 51(4), 621-634.
- Tremayne, M. (2014). Anatomy of protest in the digital era: A network analysis of Twitter and Occupy Wall Street. *Social movement studies*, 13(1), 110-126.
- Tsang, E. W. (2014). Generalising from research findings: The merits of case studies. *International Journal of Management Reviews*, 16(4), 369-383.
- Twigg, L., Taylor, J., & Mohan, J. (2010). Diversity or disadvantage? Putnam, Goodhart, ethnic heterogeneity, and collective efficacy. *Environment and planning A*, 42(6), 1421-1438.
- Twitter. (2021a). Premium Search API overview. [Online] Available at: <https://developer.twitter.com/en/docs/twitter-api/premium/search-api/overview#:~:text=Returns%20up%20to%20500%20Tweets,sent%20as%20a%20data%20body>. [Accessed 15 February 2021].
- Twitter. (2021b). Academic Research access. [Online] Available at: <https://developer.twitter.com/en/products/twitter-api/academic-research>. [Accessed 15 February 2021].
- Tyler, T. (2006). „Why do people obey the law.“ Princeton University Press. *USA and UK: Princeton and Oxford*.
- Tyler, T. R., & Fagan, J. (2008). Legitimacy and cooperation: Why do people help the police fight crime in their communities. *Ohio St. J. Crim. L.*, 6, 231.
- Udanor, C., & Anyanwu, C. C. (2019). Combating the challenges of social media hate speech in a polarised society: A Twitter ego lexalytics approach. *Data Technologies and Applications*.
- Umemoto, K. (2001). Walking in another's shoes: Epistemological challenges in participatory planning. *Journal of planning education and research*, 21(1), 17-31.
- Utz, S., & Muscanell, N. (2015). Social media and social capital: Introduction to the special issue. *Societies*, 5(2), 420-424.
- Van De Velde, B., Meijer, A. and Homburg, V. (2015) 'Police message diffusion on Twitter: analysing the reach of social media communications', *Behaviour & information technology*, 34(1), pp.4-16.

- Van der Maaten, L., & Hinton, G. (2012). Visualising non-metric similarities in multiple maps. *Machine learning*, 87, 33-55.
- Van der Meer, T., & Tolsma, J. (2014). Ethnic diversity and its effects on social cohesion. *Annual Review of Sociology*, 40(1), 459-478.
- Van Dijk, (2012). *The Network Society*. London: Sage. [ Introduction (pp. 1-8); recommended: the full Introduction (pp. 1-18)]. Available at: [www.sagepub.com/upm-data/47624\\_Van\\_Dijk.pdf](http://www.sagepub.com/upm-data/47624_Van_Dijk.pdf)
- Van Leeuwen T, (2005) *Introducing Social Semiotics*. London: Routledge.
- Van Rijsbergen, C. J. (1977). A theoretical basis for the use of co-occurrence data in information retrieval. *Journal of documentation*.
- Vance, K., Howe, W., & Dellavalle, R. P. (2009). Social internet sites as a source of public health information. *Dermatologic clinics*, 27(2), 133-136.
- Vasilopoulou, S. (2016). UK Euroscepticism and the Brexit referendum. *The Political Quarterly*, 87(2), 219-227.
- Vaterlaus, J. M., Barnett, K., Roche, C., and Young, J. A. (2016) "Snapchat is more personal": An exploratory study on Snapchat behaviors and young adult interpersonal relationships, *Computers in Human Behavior*, 62, pp. 594-601.
- Velez, M. B. (2001). The role of public social control in urban neighborhoods: A multilevel analysis of victimisation risk. *Criminology*, 39(4), 837-864.
- Wakabayashi, D. (2017) 'Google Cousin Develops Technology to Flag Toxic Online Comments', *The New York Times*, 23 February. Available at: <https://www.nytimes.com/2017/02/23/technology/jigsaw-google-toxic-comments.html> (Accessed: 13 February 2022).
- Wall, D. S., & Williams, M. (2007). Policing diversity in the digital age: Maintaining order in virtual communities. *Criminology & Criminal Justice*, 7(4), 391-415.
- Walters, M. A. (2014). Restorative approaches to working with hate crime offenders. In *Responding to Hate Crime* (pp. 243-258). Policy Press.
- Walters, M., Brown, R., & Wiedlitzka, S. (2016). Causes and motivations of hate crime. *Equality and human rights commission research report*, 102.
- Warner, B. D. (2003). The role of attenuated culture in social disorganisation theory. *Criminology*, 41(1), 73-98.

- Warner, B. D. (2007). Directly intervene or call the authorities? A study of forms of neighborhood social control within a social disorganisation framework. *Criminology*, 45(1), 99-129.
- Warner, B. D. (2014). Neighborhood factors related to the likelihood of successful informal social control efforts. *Journal of Criminal Justice*, 42(5), 421-430.
- Waterloo, S. F., Baumgartner, S. E., Peter, J., & Valkenburg, P. M. (2018). Norms of online expressions of emotion: Comparing Facebook, Twitter, Instagram, and WhatsApp. *new media & society*, 20(5), 1813-1831.
- Waters, R. D., & Jamal, J. Y. (2011). Tweet, tweet, tweet: A content analysis of nonprofit organisations' Twitter updates. *Public relations review*, 37(3), 321-324.
- Waters, R. D., Burnett, E., Lamm, A., & Lucas, J. (2009). Engaging stakeholders through social networking: How nonprofit organisations are using Facebook. *Public relations review*, 35(2), 102-106.
- Watling, S. (2011). Digital exclusion: coming out from behind closed doors. *Disability & Society*, 26(4), 491-495.
- Weaver, D. A., & Bimber, B. (2008). Finding news stories: a comparison of searches using LexisNexis and Google News. *Journalism & Mass Communication Quarterly*, 85(3), 515-530.
- Webster, F. (2001). Global challenges and national answers. *Informational societies. Understanding the third industrial revolution*, 259-278.
- Webster, F. (2014). Theories of the information society. 4th edition. London: Routledge. [chapter 1 (pp. 1-9), chapter 2 (pp. 10-33)].
- Webster, F., Blom, R., Karvonen, E., Puoskari, E., Melin, H., & Nordenstreng, K. (Eds.). (2004). *The information society reader*. Psychology Press.
- Weimann, G. (2016). Why do terrorists migrate to social media?. In *Violent Extremism Online* (pp. 45-64). Routledge.
- Weisburd, D., Davis, M., & Gill, C. (2015). Increasing collective efficacy and social capital at crime hot spots: New crime control tools for police. *Policing: A Journal of Policy and Practice*, 9(3), 265-274.
- Weitzer, R. (2003). "New Directions in Social Disorganisation Theory. *George Washington University*.

- Wells, W., Schafer, J. A., Varano, S. P., & Bynum, T. S. (2006). Neighborhood residents' production of order: The effects of collective efficacy on responses to neighborhood problems. *Crime & Delinquency*, 52(4), 523-550.
- Welsh Government (2009). *Getting On Together - a Community Cohesion Strategy for Wales*.
- Welsh Government (2016). *Respect and resilience Developing community cohesion*.
- Welsh Government (2017). *A framework for action- tackling hate crimes and incidents*.
- Wetts, R., & Willer, R. (2018). Privilege on the precipice: Perceived racial status threats lead White Americans to oppose welfare programs. *Social Forces*, 97(2), 793-822.
- Wickes, R., & Mazerolle, L. (2021). Brisbane: A story of neighbourhoods and crime in a modern Australian city. In *Crime and Disorder in Community Context* (pp. 1-22). Routledge.
- Wildemeersch, D., & Jütte, W. (2017). Digital the new normal-multiple challenges. *European Journal for Research on the Education and Learning of Adults*, 8(1), 7-20. Retrieved from [http://www.rela.ep.liu.se/issues/10.3384\\_rela.2000-7426.201681/rela\\_2017v8i1.pdf](http://www.rela.ep.liu.se/issues/10.3384_rela.2000-7426.201681/rela_2017v8i1.pdf)
- Wiles, R., Crow, G., Heath, S., & Charles, V. (2008). *Anonymity and confidentiality*.
- Williams, M. L. (2006). *Virtually criminal: Crime, deviance and regulation online*. Routledge.
- Williams, M. L. (2019). *Hatred behind the screens: A report on the rise of online hate speech*.
- Williams, M. L. (2021). *The science of hate: How prejudice becomes hate and what we can do to stop it*. Faber & Faber.
- Williams, M. L., Burnap, P. and Sloan, L (2017), Towards an ethical framework for publishing Twitter data in social research: taking into account users' views, online context and algorithmic estimation. *Sociology*
- Williams, M. L., Edwards, A., Housley, W., Burnap, P., Rana, O., Avis, N., ... & Sloan, L. (2013). Policing cyber-neighbourhoods: tension monitoring and social media networks. *Policing and society*, 23(4), 461-481.
- Williams, M. L., Sutherland, A., Roy-Chowdhury, V., Loke, T., Cullen, A., Sloan, L., ... & Giannasi, P. (2022). The effect of the Brexit vote on the variation in race

and religious hate crimes in England, Wales, Scotland and Northern Ireland. *The British Journal of Criminology*, 63(4), 1003-1023.

- Williams, M. L., Burnap, P., Javed, A., Liu, H. and Ozalp, S. (2019) 'Hate in the Machine: Anti-Black and Anti-Muslim Social Media Posts as Predictors of Offline Racially and Religiously Aggravated Crime'. *The British Journal of Criminology*.
- Williams, M.L. and Sloan, L. (2015) 'Gaining insights from social media data: collection, analysis and interpretation'.
- Williams, T. A., & Shepherd, D. A. (2017). Mixed method social network analysis: Combining inductive concept development, content analysis, and secondary data for quantitative analysis. *Organisational Research Methods*, 20(2), 268-298.
- Wong, K., & Christmann, K. (2008). The role of victim decision making in reporting of hate crimes. *Safer communities*, 7(2), 19-35.
- Wulu, J.T., Singh, K.P., Famoye, F. and McGwin, G. (2002) 'Regression analysis of count data', *Journal of the Indian Society of Agricultural Statistics*, 55(2), pp. 220-231.
- Xiao, A., Huang, Y., Bortree, D. S., & Waters, R. D. (2022). "Designing social media fundraising messages: An experimental approach to understanding how message concreteness and framing influence donation intentions." *Nonprofit and Voluntary Sector Quarterly*, 51(4), 832-856.  
doi:10.1177/08997640211022838
- Yang, J. & Counts, S. (2010). Predicting the speed, scale, and range of information diffusion in Twitter. In *Proceedings of the Fourth International Conference on Weblogs and Social Media* (p. 4). Association for the Advancement of Artificial Intelligence.
- Yoo, B., & Rayz, J. T. (2021). Understanding emojis for sentiment analysis. In *The International FLAIRS Conference Proceedings* (Vol. 34).
- Yuan, X., & Crooks, A. T. (2018). Examining online vaccination discussion and communities in Twitter. In *Proceedings of the 9th international conference on social media and society* (pp. 197-206).
- Zappavigna, M. (2011) 'Ambient affiliation: a linguistic perspective on Twitter', *New Media & Society*, 13(5), pp.788-806.
- Zarrella, D. (2009), *The Science of Retweets*, available online at <http://danzarrella.com/science-of-retweets.pdf>.

Zedner, L. (2004), *Criminal Justice*, Oxford: Oxford University Press.

Zhang, L., Messner, S. F., & Liu, J. (2007). A multilevel analysis of the risk of household burglary in the city of Tianjin, China. *The British Journal of Criminology*, 47(6), 918-937.

Zhang, S., de Jong, M. D., & Gosselt, J. F. (2022). Microblogging for engagement: Effects of prior company involvement, communication strategy, and emojis on Western and Chinese users. *Journal of International Consumer Marketing*, 34(5), 616-630

Zhou, X. and Tu, W. (2000) 'Confidence intervals for the mean of diagnostic test charge data containing zeros', *Biometrics*, 56(4), pp. 1118-1125.

Ziems, C., He, B., Soni, S., & Kumar, S. (2020). *Racism is a virus: Anti-asian hate and counterhate in social media during the covid-19 crisis*. Georgia Institute of Technology.

# Appendices

---

## Appendix 1: Survey (Chapter One)

# Cohesion Delivery Network Survey

---

Start of Block: Default Question Block

Only sign here if you did not sign the provided consent form prior to the survey

---

Q1 Which of the following organisations or roles do you best represent?

- Cohesion Coordinators (1)
  - Cohesion Officers (2)
  - The Welsh Government (3)
  - Community Groups (4)
  - Academic (5)
  - The Police (6)
- 

Q2 Which police region do you operate in?

- Dyfed-Powys Police (1)
  - Gwent Police (2)
  - North Wales Police (3)
  - South Wales Police (4)
  - Across two or more regions (5)
-

Q3 What is your specific role in your organisation?

---

---

Q4 How long have you been in your post for?

- Less than 3 months (1)
- Between 3 and 12 months (2)
- Between 1 and 2 years (3)
- Between 2 and 5 years (4)
- Over 5 years (5)

---

Q7 Has your role received any funding from the EU Transition Fund?

- Yes (1)
- No (2)
- Dont know (3)

---

Q8 Is your role government based?

- Yes (1)
- No (2)
- Dont know (3)

---



Q5 How important is cooperation with other agencies for the delivery of cohesion?

- Extremely important (1)
  - Very important (2)
  - Moderately important (3)
  - Slightly important (4)
  - Not at all important (5)
- 

Q6 How important do you feel it is to improve cooperation within the network?

- Extremely important (1)
  - Very important (2)
  - Moderately important (3)
  - Slightly important (4)
  - Not at all important (5)
-

Q13 How much does your organisation focus on the following protected characteristics (The Equality Act)?

	A great deal (1)	A lot (2)	A moderate amount (3)	A little (4)	None at all (5)
Age (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Disability (2)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Gender Reassignment (3)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Marriage and Civil Partnership (4)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pregnancy and Maternity (5)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Race (6)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Religion or Belief (7)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sex (8)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sexual Orientation (9)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

---

Q14 How effective would you rate your organisation in the following outcomes:

	Extremely effective (1)	Very effective (2)	Moderately effective (3)	Slightly effective (4)	Not effective at all (5)
Preventing tensions (6)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Detecting tensions (7)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Responding to tensions (8)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

---

Q15 From your organisation's perception will tensions roughly rise, fall or stay the same over the next 5 years?

- Rise (1)
- Fall (2)
- Stay the same (3)

---

Q16 From your organisation's perception will Hate Crimes roughly rise, fall or stay the same over the next 5 years?

- Rise (1)
- Fall (2)
- Stay the same (3)

---

Q17 For this question you will be shown each agency being studied for this research. For each you must indicate the degree to which cooperation frequency exists between yourself and the given agency. The criteria indicators for this comparison task are given below:

1 = Very regular cooperation (several times a week) 2= Regular cooperation (every week) 3= Some cooperation (every few weeks) 4= Sporadic cooperation (cooperate, but not a frequent level) 5 = Minimum cooperation (have cooperated but very

infrequent) 6= No cooperation N/A= If this is the agency you are a part of  
 It is important you consider your response carefully and try and give as accurate information as possible that fits within the given criteria.

	1 (1)	2 (2)	3 (3)	4 (4)	5 (5)	6 (6)	N/A (7)
Cohesion Coordinators (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cohesion Officers (2)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Welsh Government (3)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Community Group (4)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Academic (5)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The Police (6)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q18 In order to enhance the network would you cooperate more, less of the same with the following organisations?

	More (1)	Less (2)	Same (3)
Cohesion Coordinators (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cohesion Officers (2)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Welsh Government (3)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Community Group (4)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Academic (5)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The Police (6)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q19 What are the reasons for your current level of cooperation frequency with other stakeholders? If you plan to change the current level of cooperation frequency, please provide the reasons why.

---



---

Q20 For this question you will again be shown each agency being studied for this research. For each you must indicate your perceived quality of cooperation between yourself and the given agency. No criteria indicators exist for this comparison task, this is based on your own insight and perception. However, the following is given to indicate the direction of the scale:

1 = Excellent quality of cooperation 6= Poor quality of cooperation

	1 (1)	2 (2)	3 (3)	4 (4)	5 (5)	6 (6)	N/A (7)
Cohesion Coordinators (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cohesion Officers (2)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Welsh Government (3)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Community Group (4)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Academic (5)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The Police (6)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q21 Please elaborate on reasons why you gave specific scores the previous question

---



---

Q22 How much responsibility do you think the following organisations should have for 'cohesion delivery' in Wales?

	A great deal (1)	A lot (2)	A moderate amount (3)	A little (4)	None at all (5)
Cohesion Coordinators (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cohesion Officers (2)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Welsh Government (3)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Community Group (4)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Academic (5)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The Police (6)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q9 How much would you say COVID-19 has impacted your work?

- A great deal (1)
- A lot (2)
- A moderate amount (3)
- A little (4)
- None at all (5)

Q10 In what ways has COVID-19 impacted your work?

---



---

Q11 How much has COVID-19 impacted your cooperation with other stakeholders responsible for the delivery of cohesion in Wales?

- A great deal (1)
  - A lot (2)
  - A moderate amount (3)
  - A little (4)
  - None at all (5)
- 

Q12 In what ways has COVID-19 impacted your cooperation with other stakeholders responsible for the delivery of cohesion in Wales?

---

---

Q23 How much do you agree with the following statement:

Further methods of online tension monitoring should be implemented within the network

- A great deal (1)
  - A lot (2)
  - A moderate amount (3)
  - A little (4)
  - None at all (5)
- 

Q24 Please give a brief explanation for your previous selection

---

End of Block: Default Question Block

---

## **Appendix 2: Information Sheet**

# **Welcome to the Cohesion Network Mapping Survey**

This survey will give a valuable contribution to a greater research project that seeks to examine the delivery of cohesion in Wales. This project will assess both offline and online methods used and seek to create recommendations for the network moving forward. The research is funded by...

**Who should complete the survey?** Organisations and individuals that share responsibility for the delivery of community cohesion in Wales. The main cluster groups identified are 1) Cohesion Coordinators 2) Cohesion Officers 3) The Welsh Government 4) Community Groups 5) Charities 6) Academic 7) The Police. Participants should have a knowledge of tension and hate crime issues in Wales and the methods used to prevent, identify or respond to them.

**The survey should take 15-20 minutes to complete.**

**Context of the study** Recent years has presented new challenges and developments within cohesion delivery networks within the United Kingdom. Most significantly has been the implications of Britain leaving the European Union and the development of online tensions and hate crimes. These two factors have given way to a required change in cohesion delivery. For example, the European Transition Fund has led to new roles being created within the cohesion network. Furthermore, this study seeks to understand the current network and additionally examine what impact changes could have to it. This particular survey seeks to map out the dynamic of various relationships within the current network and find potential gaps and bottlenecks.

### **Anonymity, Confidentiality and Data Protection**

All participants will remain anonymous throughout the process and in the final output. However, some personal data that can identify particular respondents will be collected but will only be accessed by Ben Foster (the researcher). This is collected to check for duplicate responses. The data collected will only be used for the purposes outlined previously. Once collected, the data will be stored in either a



secure Cardiff University Network computer or encrypted laptop. All data will be retained in accordance with the Data Protection Act (1998) until six months after the completion of the thesis when it will all be destroyed. Cookies etc. are not used in this survey.

If you have any further questions regarding the above or the research in general, please contact:

**Ben Foster:** [fosterb1@cardiff.ac.uk](mailto:fosterb1@cardiff.ac.uk)

## Appendix 3: Consent Form



### Consent Form

#### Cohesion Network Mapping Survey

Dear Participant,

I am organising a survey that seeks to generate data appropriate for mapping out the cohesion delivery network in Wales. This survey seeks to examine lived experiences of stakeholders in the network and understand the dynamic of their relationships across agencies. Further information about the nature of the survey and how the data will be used is included in the attached information sheet.

If you are able and willing to be a participant in this research, please complete the attached online survey; 'Cohesion Network Mapping Survey'. The survey typically takes between 15 and 20 minutes. The questions are typically presented through likert scales giving multiple choices. There are, however, some open questions that will enable you to explain particular perspectives you have about the network. The survey seeks to help improve the network you work within therefore detailed responses are welcomed.

What will happen to the information gathered in this research?

The research will be written up as part of a PhD thesis. Your name will NOT be used at any point in the final thesis. Additionally, you will remain fully anonymous, as any identifiable information collected will be carefully removed. If you agree to participate in the study you are able to choose to withdraw at any given time.

I have been fully informed by the researcher of the aims of this research.

I understand that:

- am not required to participate in this research project and, if I choose to participate, I can withdraw from taking part at any stage in the process.

- have the right to refuse permission for the publication of any information about me
- Any information that is collected will be only for the purpose of this research project.
- All of the information I give will be treated as strictly confidential
- The researcher will make every effort to ensure that anonymity of the participant is preserved.

.....

(Signature of participant)

.....

(Date)

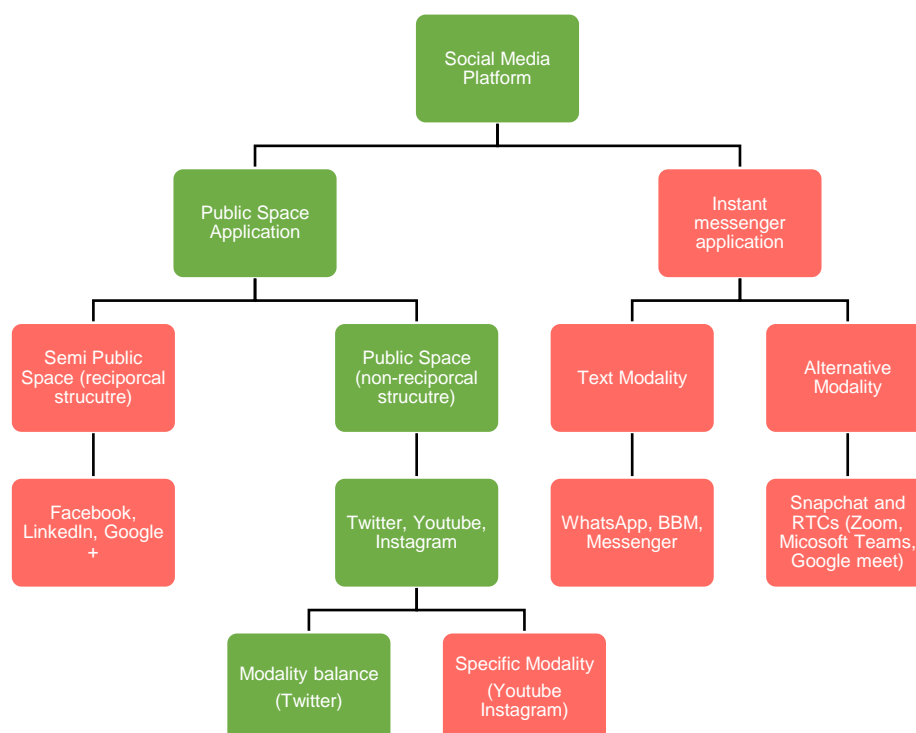
.....

(Printed name of participant)

## Appendix 4: Social Media Platform Selection

In determining which Web 2.0 function is most appropriate to measure for communication strategies, it was necessary to consider the various sources of Web 2.0 communications data, which can vary significantly (Burnap et al., 2015). These sources include RSS feeds, personal blogs, wikis, and microblog social networking platforms, such as Twitter, Instagram, Facebook, TikTok, and MySpace. Given the focus on communication strategies in this study, social networking platforms were deemed the most suitable source.

Appendix 3 details fundamental differences in social media platforms. Based on such differences, it outlines the decision-making process for selecting the most relevant platform for research. Platforms can be categorized into instant messenger applications and public space applications (Waterloo, 2018). Instant messenger applications allow communication through private channels, which may be limited to specific groups. In contrast, public space applications must operate in public spaces to some extent (Karapanos et al., 2016). To comprehend how third sector organizations communicate with digital audiences, public space applications were deemed the most suitable option for this analysis.



Public space applications can vary based on their behavioural privacy settings, which determine the level of public accessibility of content (Lapinski and Rimal, 2005). Davenport et al. (2014) categorizes these settings into two types: reciprocal and non-reciprocal. Reciprocal or closed friendship networks occur when users have agreed to connect online, while non-reciprocal or open friendship networks do not require prior connections (Lup et al., 2015). Reciprocal structures have limited access to public discourse, whereas non-reciprocal structures allow for open discourse and have been compared to a digital agora or public square (Burnap et al., 2015).

## Appendix 5: Google Perspective API/ Commanalytic Set Up and Tutorial

*Ben Foster*  
*Cardiff University*

Tools: Google Perspective API, Commanalytics

Permissions: Google Perspective API

### **Step 1:** Gaining Google Perspective API Approval

- 1) Visit: <https://www.perspectiveapi.com/>
- 2) Click 'get started' (top right)
- 3) Choose 'developers'
- 4) Follow
- 5) Click 'get started' (pink box)
- 6) Follow instructions:

Prerequisites:

- I. You must have a Google account, giving you access to the suite of Google products including Google Cloud.
- II. You also must have a Google Cloud project to authenticate (but not necessarily host) your API requests. Go to the Google Cloud console and use an existing project or follow these steps to create a new one:
  - 7) Sign into google account (click 'go to console')
  - 8) On console page either create 'new project' or 'select project' (screen should look like extract below)

- 9) Name the project: “Note the project ID, shown under the Project name field. The project ID will be used to identify your project for Google Cloud tools”
- 10) Click ‘create’, project will now appear in drop down menu
- 11) Request API approval, by clicking ‘request API access’ or follow this link: <https://docs.google.com/forms/d/e/1FAIpQLSdhBBnVVVbXSElby-jhNnEj-Zwpt5toQSCFsJerGfpXW66CuQ/viewform>

Once all of the previous steps have been completed google perspective API access will be enabled on your given google account 👍 👍 👍

However, google perspective API requires another tool to actually be used:

### ‘Communalytics’

Communalytics Edu requirement: academic email address i.e. @cardiff.ac.uk

*Communalytics Edu Features:*

- **Account Caps:** ≤ 30K records shared across 3 datasets
- **Simultaneously Run Multiple Data Collectors:** 1 Reddit, 1 Twitter & 1 CrowdTangle
- **Collaboration Friendly:** Access to various team collaboration features
- **Toxicity Analysis:** Pinpoint toxic and anti-social interactions in your dataset w/Google Perspective API

- **Sentiment Analysis:** Detect the polarity (neutral, negative, or positive) of posts in your dataset
- **Network Visualizer:** Discover and build various types of networks via the built-in network analysis and visualization module
- **Twitter Bot Detection:** Detect the presence of Twitter bot accounts and the use of automation w/Botometer API

#### Data Sources:

- **Reddit:** Historical data collection for  $\leq 7$  consecutive days
- **Crowdtangle Facebook/Instagram URL Search:** Public Facebook/Instagram posts ( $\leq 10k$ ) that shared the same URL(s) (Req. CrowdTangle Account)
- **Twitter Threads:** Public replies ( $\leq 10K$ ) to any public tweet posted within the previous 7 days (Req. a Twitter Developers Account)
- **Twitter Academic Research Product Track:** Not supported
- **Import/Rehydrate Twitter IDs:** Ability to recollect tweets based on a list of unique Twitter IDs
- **Import CSV Files:** Ability to import an existing dataset from a CSV file
- **Import Twitter JSON Files:** Not supported

#### Step 2: Communalytics set up

- 1) Visit <https://communalytic.com/>
- 2) Scroll down the opening page to the sign up boxes:



- 3) Select the box on the left (sign up now for free edu account)
- 4) **Confirm online identity:** Use a **GMAIL\*** account to confirm your online identity. ( We're unable to support other email accounts at this time.) **\*Note:** This is the



email you will use to login into **Communalytic EDU** once your account is approved.

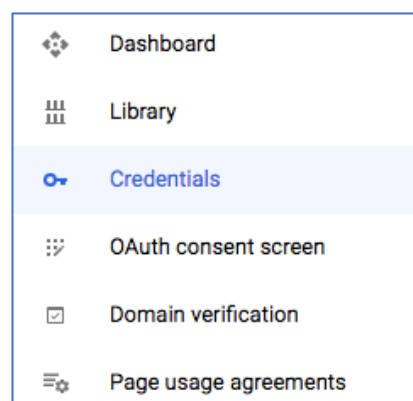
- 5) **Confirm Academic Affiliation: Communalytic EDU** is only available for academic research, we will need to send a confirmation email to your academic email address\* (e.g., name@syr.edu, [name@ryerson.ca](mailto:name@ryerson.ca)). **\*Note:** If your institution uses GMAIL for email services, you can use the same address in Step 1 & 2.

(If academic email is not recognised: “We use a comprehensive list of academic international domain names to confirm your academic email from Step 2, but if it’s not properly recognized by our system, we’ll ask you to provide more details about your academic affiliation via a short online form”)

### **Step 3)** Linking Google Perspective API to Communalytics Dashboard


Once both step 1 and step 2 are complete, you should have access to google perspective api and a Communalytics (edu) account, however they will not be linked (meaning you cant use the google perspective api on the Communalytics dashboard). To enable google perspective api on Communalytics follow the next steps:

- 1) Return to the google console page (where you created project in step 1)
- 2) Select ‘credentials’ in the left hand bar seen below:



- 3) Provided that you have successfully gained access to persepcitve API, the next window should show the following:

<input type="checkbox"/>	Name	Creation date ↓	Restrictions	Key	Actions
<input type="checkbox"/>	API key 1	Sep 22, 2021	None	[Redacted]	[Edit] [Delete]

- 4) Copy the key provided in the box
- 5) Return to your logged in Communalitics page
- 6) Click this icon:  (top right)
- 7) Select 'my profile'
- 8) Scroll down to the 'API Keys' box
- 9) By 'perspective api' select the enter key box (green)
- 10) Paste the key you copied earlier (step 3.4)
- 11) Select: Submit

Once these steps have been completed you should now be able to use Google perspectives api on the Communalitics dashboard

#### Step 4: Using GPAPI on Communalitics

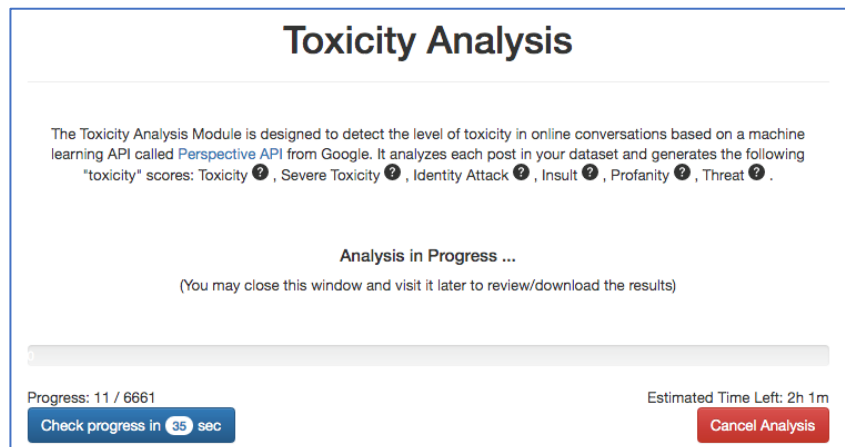
Many different data sources can be used in Communalitics, if you are importing a CSV file with twitter data externally collected (e.g. on COSMSOS) you will need to follow these steps:

- 1) Click 'my datasets (top right)
- 2) Select CSV file
- 3) Format the CSV file using the template provided- this step is very important as the requirements must be met in order for the file to be processed

Requirements:

- I. Follow template provided
  - II. Must include following columns: **(1) created\_at, (2) user\_screen\_name, (3) text and (4) in\_reply\_to\_screen\_name**
- 4) Once dataset is prepared, click 'choose file', and select dataset

- 5) Select Import CSV file
- 6) The dataset will now exist at the bottom of the 'my dataset' page
- 7) Under 'dataset name' click on the link (dataset title)
- 8) On the dashboard (left of screen) scroll down to and click on 'toxicity analysis'
- 9) In new window select 'start analysis'
- 10) The toxicity analysis for each tweet will begin (as seen below), this can take some time, so leave the web browser open



At the end of the process, the toxicity scores will be available in new columns. The new dataset can be downloaded by clicking 'download dataset' in the bottom right screen.

Further tutorials for Communalitics: <https://communalitics.com/video-tutorials/>  
About google perspective API: <https://developers.perspectiveapi.com/s/about-the-api>  
Examples of GPAPI in use: <https://www.perspectiveapi.com/case-studies/>

References if used in research:

Gruzd, A., & Mai, P. (2021). Communalitics: A Research Tool For Studying Online Communities and Online Discourse. Available at <https://Communalitics.com>

Ben Foster

Cardiff University, School of Social Sciences

## Appendix 6: Code Manual (Chapter 6)

### Code Manual – Third Sector Organisation Social Media Strategies

#### Overview of project

The aim of this coding process is to categorise all tweets in the dataset, first into Mergel's (2013) broad communication strategies (push, pull, networking and transactional). During the initial coding phase other 'content factors' will be measured, such as if a tweet mentions covid or other protected characteristics. The second round of coding helps in making sense of differing approaches within some of the broad strategies. This is applied to push and pull strategies, creating eight new 'sub-strategies' that were thematically identified during the first phase in coding. Networking and transactional cases were deemed relatively cohesive and were therefore not further sub-categorised.

#### Automatic Coding/ Variables

These factors were automatically captured during the data collection phase

- 1) Tweet Date
- 2) Tweet Type
- 3) Organisation Name
- 4) Tweet Text (String)
- 5) Hashtag (String)
- 6) Mentions (String)
- 7) Likes
- 8) Replies
- 9) Retweets
- 10) Hyperlink Used

#### Quick Coding/ Variables

These factors are quickly established using quick functions on SPSS

- 1) Hashtag Count (word frequency of hashtag string)
- 2) Mention Count (word frequency of mention string)
- 3) Word Count (word frequency of tweet text string)

### 1.1 [Account Size]

**How big is the account?** Based on account followers

<b>1</b>	<b>Small</b>	Account has: 0-3000 followers
<b>2</b>	<b>Medium</b>	Account has: 3000-6000 followers
<b>3</b>	<b>Large</b>	Account has: 6000-10000 followers
<b>4</b>	<b>X-Large</b>	Account has: 10000+ followers

### 1.2 [Tweet Frequency]

**What is the account's tweet frequency?** C

<b>1</b>	<b>Low</b>	0-1.5
<b>2</b>	<b>Mid</b>	1.5-2.5
<b>3</b>	<b>High</b>	2.5+

### 1.3 [Time of Day]

**What time of day was the tweet authored?** (each time of day should be further coded into a dummy variable)

<b>1</b>	<b>Morning</b>	Tweet time: 6:00 am – 11:00 am
<b>2</b>	<b>Daytime</b>	Tweet time: 11:00 am – 5:00 pm
<b>3</b>	<b>Night</b>	Tweet time: 5:00 pm – 10:00 pm
<b>4</b>	<b>Evening</b>	Tweet time: 10:00 pm – 6:00 am

#### 1.4 [Weekend]

Was the tweet authored on the weekend?

0	No
1	Yes

#### 1.5 [Hashtag Type]





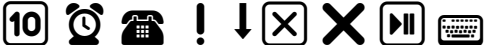
What hashtag type (as defined by Saxton et al., 2015) is used ? (each time of day should be further coded into a dummy variable). Note: only apply to a hashtag used four or more times in the dataset

1	<b>Knowledge/ Informative</b>	Any hashtag that seeks to help inform audiences on concepts and terms associated with the organisation's work
2	<b>Events</b>	Events that are either organisation run, or a global day/ week of recognition
3	<b>Call-To-Action</b>	Hashtags encouraging the audience to do something
4	<b>Values and Goals</b>	Organisational values or goals. Useful for reinforcing the organization's core values and ultimate strategic goals
5	<b>Organisation Branding</b>	Organisation-specific hashtags, unique organization identifiers, hashtags noting one of the organization's programs names
6	<b>External Branding</b>	The same criteria as organisation branding, but for an organisation or movement autonomous to the author organisation
7	<b>Time and Place</b>	Hashtag relating to time or location
8	<b>Dialogical</b>	

Use ISNUMBER function on excel

## 1.6 [Emoji Type]

**What emoji type is used ?** (each time of day should be further coded into a dummy variable).

1	<b>Emotive</b>	An emoji used to emphasise and/ or express an emotion of the author	
2	<b>Group Specific</b>	An emoji that references a particular social group or movement e.g., 'rainbow flag' for LGBT, 'black fist' for BLM, 'female sign' for women/ feminism	 
3	<b>Activity/ Weather</b>	An emoji that promotes an activity, sometimes using weather, discussed in any given tweet. These are often used when promoting an organisation-based activity.	
4	<b>Symbolic</b>		

Use ISNUMBER function on excel

## Coding

### Round 1 Features

#### 2.1 [Broad Communication Strategy]

What is the MAIN purpose of this post? Code only one. Strategies reflect Mergel's (2013) social media communication strategy typology

1	<b>Push</b>	Any tweet that: <b>broadcasts information</b> on the platform with <b>no direct</b> aim of <b>instigating conversation</b>
2	<b>Pull</b>	Any tweet that: attempts to redirect audiences to other <b>organisation affiliated</b> content that is <b>not on Twitter</b>
3	<b>Networking</b>	Any tweet that: attempts to <b>elicit donations</b> or promote <b>organisation-run services</b>
4	<b>Transactional</b>	Any tweet that: attempts to <b>engage followers</b> in <b>on-platform discussion</b>

2.2 [Other Focus] Does the post reference or provide support for another protected characteristic, that is not the primary focus of the author organisation

0	<b>No</b>
1	<b>Yes</b>

2.3 [Mention Covid] Does the post reference Covid 19. Trigger words include: 'covid', 'corona', 'covid19', 'lockdown', 'corona virus', 'c19' and 'vaccine'

0	<b>No</b>
1	<b>Yes</b>



2.3 **[Celebration ]** Does the post reference a day, week or month of celebration?.

0	No
1	Yes

2.3 **[Intervention ]** Any post that publicly condemns a specific instance of hate, which could refer to both online and offline incidents

0	No
1	Yes

### Round 2 Features

3.1.1 **[Push Sub-Strategies]**

What features of this post are most in line with the following options? If the initial round of coding did not identify the post as ‘push strategy’ skip this step

1	<b>Non-Organisation Information</b>	Any tweet that discusses concepts that do not directly relate to the organisation’s operations.
2	<b>Link to other content (non-social media)</b>	Any tweet that disseminates information by providing links to non-organization webpages, such as news articles
3	<b>Link to other content (social media)</b>	Any tweet that uses the @ symbol to mention other nodes while broadcasting information
4	<b>Organisation Information</b>	Any tweet that mentions the work or operations carried out by the organization

### 3.1.2 [Push Sub-Strategy Selection]

For the selected sub-strategy, code it (within its unique variable) 'yes' (1)

<b>0</b>	<b>No</b>
<b>1</b>	<b>Yes</b>

### 3.2 [Pull Sub-Strategies]

**What features of this post are most in line with the following options?** If the initial round of coding did not identify the post as '**pull strategy**' skip this step

<b>1</b>	<b>Interactive</b>	Any tweet that appeals to audiences to participate in organization-run activities, such as focus groups or surveys
<b>2</b>	<b>Work Recruitment</b>	Any tweet that attempts to recruit new staff or volunteers through tweets
<b>3</b>	<b>Observational</b>	Any tweet that encourages followers to follow a hyperlink to websites belonging to the third sector organization that published the tweet.

### 3.2.2 [Pull Sub-Strategy Selection]

For the selected sub-strategy, code it (within its unique variable) 'yes' (1)

<b>0</b>	<b>No</b>
<b>1</b>	<b>Yes</b>

### 3.3 [Transactional Sub-Strategies]

What features of this post are most in line with the following options? If the initial round of coding did not identify the post as 'pull strategy' skip this step

<b>1</b>	<b>Elicit Donations</b>	Any tweet that attempts to elicit donations or fundraise for the organisation
<b>2</b>	<b>Service Use</b>	Any tweet that promotes an <u>organisation-run</u> service

#### 3.3.2 [Transactional Sub-Strategy Selection]

For the selected sub-strategy, code it (within its unique variable) 'yes' (1)

<b>0</b>	<b>No</b>
<b>1</b>	<b>Yes</b>

## Appendix 7: Ethical Approval



School of  
Social Sciences  
Ysgol y Gwyddorau  
Cymdeithasol

**Cardiff University**  
Glamorgan Building  
King Edwards VII Avenue, Cardiff  
CF10 3WT, Wales, UK  
Tel +44(0)29 2087 5179  
Fax +44(0)29 2087 4175  
[www.cardiff.ac.uk/social-sciences](http://www.cardiff.ac.uk/social-sciences)

29 May 2019

Our ref: SREC/3268

Benjamin Foster  
SOCSI

**Prifysgol Caerdydd**  
Adelad Morgannwg  
Rhodfa'r Brenin Edward VII, Caerdydd  
Ffôn +44(0)29 2087 5179  
Ffacs +44(0)29 2087 4175  
[www.cardiff.ac.uk/social-sciences](http://www.cardiff.ac.uk/social-sciences)

Dear Benjamin,

Your project entitled *'Investigating Social Media and Offline Data to Measure Extremism and Hate Crime'* has now been approved by the School of Social Sciences Research Ethics Committee of Cardiff University and you can now commence the project should all necessary forms of approval been received.

If you make any substantial changes with ethical implications to the project as it progresses you need to inform the SREC about the nature of these changes. Such changes could be: 1) changes in the type of participants recruited (e.g. inclusion of a group of potentially vulnerable participants), 2) changes to questionnaires, interview guides etc. (e.g. including new questions on sensitive issues), 3) changes to the way data are handled (e.g. sharing of non-anonymised data with other researchers).

In addition, if anything occurs in your project from which you think the SREC might usefully learn, then please do share this information with us.

All ongoing projects will be monitored and you will be obliged periodically to complete and return a SREC monitoring form.

Please inform the SREC when the project has ended.

Please use the SREC's project reference number above in any future correspondence.

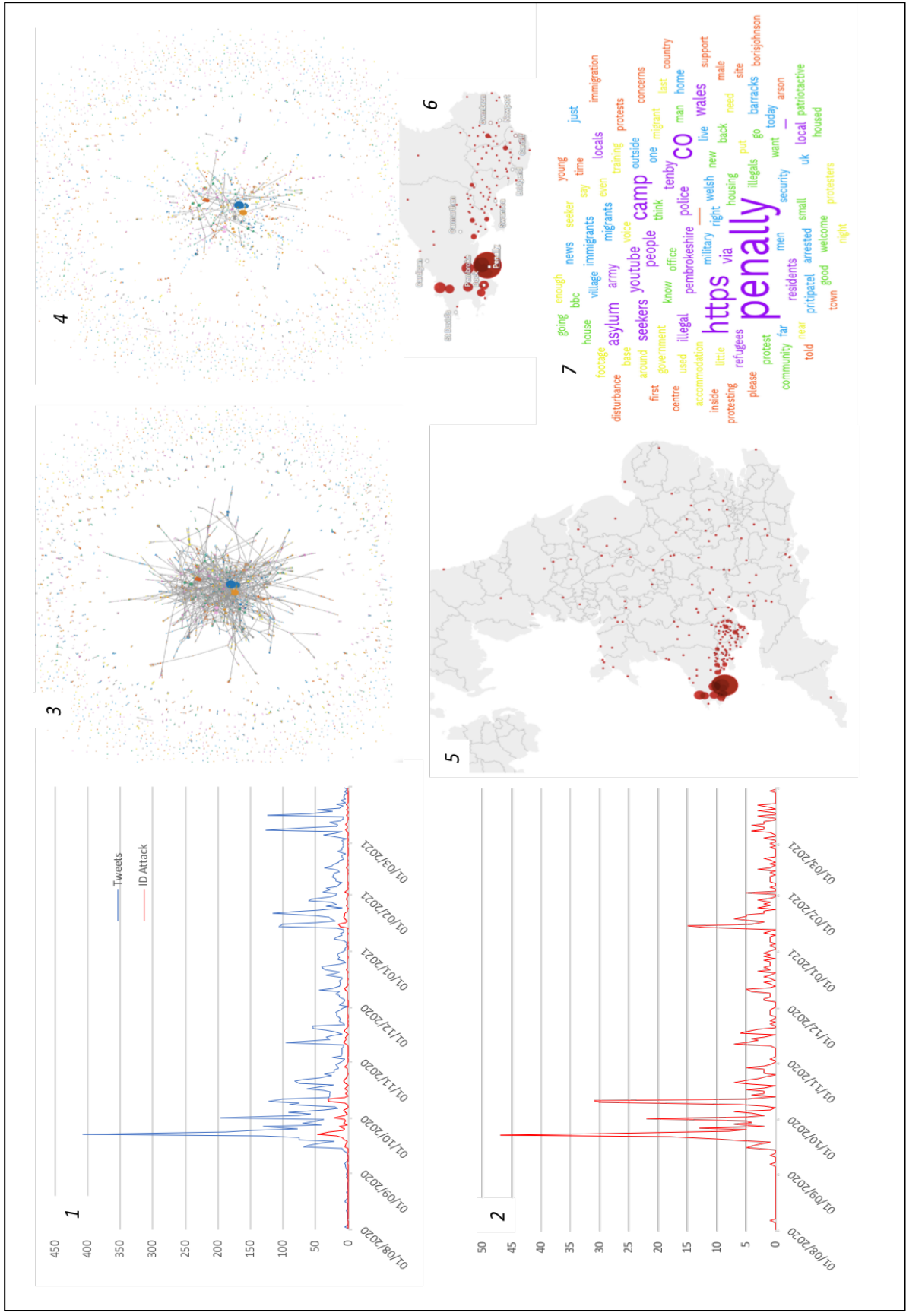
Yours sincerely

**Professor Alison Bullock**  
**Chair of School of Social Sciences Research Ethics Committee**  
Cc Matthew Williams



Registered Charity No. 1136855  
Elusen Gofrestredig Rhif. 1136855

# Appendix 8: Data Visualisation (Dashboard Format)



## Appendix 9: PCA Results (Chapter Five)

	Content Factors	Social Factors	External Factors	Controls
<b>N</b>	<b>6,661</b>	<b>6,661</b>	<b>6,661</b>	<b>6,661</b>
<b>-2 LL</b>	<b>-18787.98</b>	<b>-20085.00</b>	<b>-20479.56</b>	<b>-20515.15</b>
<b>BIC</b>	<b>37646.34</b>	<b>40258.04</b>	<b>40985.54</b>	<b>41100.72</b>

### Sub-factor analysis for size model (likes)

	Content Factors	Social Factors	External Factors	Controls
<b>N</b>	<b>6,661</b>	<b>6,661</b>	<b>6,661</b>	<b>6,661</b>
<b>-2 LL</b>	<b>-12888.53</b>	<b>-14229.68</b>	<b>-14810.59</b>	<b>-14838.31</b>
<b>BIC</b>	<b>25847.44</b>	<b>28547.40</b>	<b>29647.59</b>	<b>29755.06</b>

### Sub-factor analysis for size model (retweets)

	Content Factors	Social Factors	External Factors	Controls
<b>N</b>	<b>6,661</b>	<b>6,661</b>	<b>6,661</b>	<b>6,661</b>
<b>-2 LL</b>	<b>-10215.39</b>	<b>-11592.59</b>	<b>-11724.47</b>	<b>-11386.93</b>
<b>BIC</b>	<b>20501.15</b>	<b>23273.23</b>	<b>23475.35</b>	<b>22844.29</b>

### Sub-factor analysis for size model (replies)

## Appendix 10: PCA Results (Chapter Six)

	Full Model	Strategy Factors	Content Factors	Account Factors	Controls (Time)	Hashtag/Emoji Type
<b>N</b>	<b>5000</b>	<b>5000</b>	<b>5000</b>	<b>5000</b>	<b>5000</b>	<b>5000</b>
<b>-2 LL</b>	<b>-13942.26</b>	<b>-14546.01</b>	<b>-14383.32</b>	<b>-14580.22</b>	<b>-14806.37</b>	<b>-14741.65</b>
<b>BIC</b>	<b>28148.54</b>	<b>29126.09</b>	<b>28843.29</b>	<b>29211.55</b>	<b>29655.31</b>	<b>29585.50</b>

### Sub-factor analysis 'likes' model

	Full Model	Strategy Factors	Content Factors	Account Factors	Controls (Time)	Hashtag/Emoji Type
<b>N</b>	<b>5000</b>	<b>5000</b>	<b>5000</b>	<b>5000</b>	<b>5000</b>	<b>5000</b>
<b>-2 LL</b>	<b>-10936.13</b>	<b>-11734.72</b>	<b>-11419.49</b>	<b>-11914.83</b>	<b>-12193.91</b>	<b>-12133.04</b>
<b>BIC</b>	<b>22136.28</b>	<b>23503.51</b>	<b>22915.63</b>	<b>23880.76</b>	<b>24430.41</b>	<b>24385.31</b>

### Sub-factor analysis 'retweets' model

	Full Model	Strategy Factors	Content Factors	Account Factors	Controls (Time)	Hashtag/Emoji Type
<b>N</b>	<b>5000</b>	<b>5000</b>	<b>5000</b>	<b>5000</b>	<b>5000</b>	<b>5000</b>
<b>-2 LL</b>	<b>-3051.83</b>	<b>-3295.87</b>	<b>-3173.12</b>	<b>-3175.56</b>	<b>-3276.46</b>	<b>-3266.50</b>
<b>BIC</b>	<b>6367.69</b>	<b>6625.81</b>	<b>6422.88</b>	<b>6402.22</b>	<b>6595.50</b>	<b>6652.24</b>

### Sub-factor analysis 'replies' model

## Appendix 11: Twitter Scraping Selection Processes

(For Chapter Six)

With four distinct options identified by the researcher, a list of five requirements was established to assess which was the best fit, in accordance with the pragmatist approach, to determine the "what works" approach for Twitter collection (Creswell and Plano Clark, 2018). The first requirement was to collect the tweet string data, which was essential for both the quantitative and qualitative aspects of the study. The tweet string data was necessary for manual coding processes to allow for content analyses that included both approaches, which will be further discussed in the section.

The second requirement was to collect all tweets, not a sample. The third was to include engagement data, such as likes, retweets, and replies (Boyd et al., 2010), which is crucial to determine the most endorsed strategies and content factors (Williams and Burnap, 2016; Ozlap et al., 2020). This will be discussed in more detail when analysing the data. The fourth requirement was to collect historic data to understand day-to-day tolerance promotion strategies, rather than those in response to a specific event, as most other studies have done (Earl et al., 2013; Tremayne, 2014).

The COSMOS tool, which only collects data from the previous 14 days, was not considered (Burnap et al., 2015). Data extraction from a specific account, instead of a generalised stream of multiple users, was essential for exploring third-sector organizations on Twitter. The ease of access to the Twitter API application process was the final requirement. Based on a detailed consideration of these six requirements, a pragmatic extraction method was chosen that best fit the study's aims.

Requirement:	N-Capture	Cosmos	Search API	Streaming API
Collects Tweet String	✓	✓	✓	✓
Census of Tweets	✓	✗	✓	✗
Engagement Data	✓	✓	✓	✓
Historic Data	✓	✗	✓	✗
Single Account Collection	✓	✗	✓	✗
Easy Access	✓	✓	✗	✗

Due to the extremely high volumes of tweets in these cases, samples of tweets relating to the trigger events were collected, rather than an overall census. Thereafter, in both cases extraction tools such as COSMOS, were preferred, because they implement a persistent connection to Twitter 1% stream (Williams and Sloan, 2015). While using sampling is necessary for vast collections in wide-scale trigger events such as these, the Penally case study is significantly more localised. This presents a unique opportunity to collect a near-census<sup>77</sup> of all tweets, providing a truly representative dataset.

(For Chapter Five)

To allow for the comparison of cyberhate levels before, during, and after trigger events, historical data collection was preferred, which aligned with prior research. However, not using live collection tools like COSMOS can result in missing relevant online content that may be removed or deleted post-trigger events. Besides, geotag data is present in only 1% of tweets, making its inclusion less critical (Sloan et al., 2013). Thus, geotag functions were not prioritized when selecting the data extraction method, and alternative approaches for analysing locational changes in informational flows are discussed later.

Requirement:	N-Capture	Cosmos	Search API	Streaming API	Academic
Collects Tweet String	✓	✓	✓	✓	✓
Census of Tweets	✓	×	✓	×	✓
Engagement Data	×	✓	✓	✓	✓
Keyword Search	×	✓	✓	✓	✓
Complete Historic Data	✓	×	✓	×	✓
No Collection Limit	×	✓	×	✓	✓

In addition to the criteria used in the previous section, some additional criteria were implemented for this study. Instead of collecting tweets authored by pre-selected accounts, as seen in Chapter Six, a keyword search term was used to provide data from a wide range of social actors, creating what is referred to as a "digital agora" (Williams and Burnap, 2016: 218). Therefore, the tool used for data collection required

---

<sup>77</sup> A true census was not achieved as tweets about Penally may exist outside of the keyword Penally, and were therefore not collected



keyword-based collections. Another difference in the criteria was that there were no restrictions on the data collection limit.

Unlike Chapter Six, which had a pre-determined tweet collection limit, this chapter conducted a census of all tweets related to a specified keyword search within a specific timeframe. This approach required an extraction tool without a tweet collection limit. After identifying the criteria, five extraction tools were evaluated. NCapture lacked the ability to collect engagement data and had an inadequate keyword search function. COSMOS and Streaming API couldn't collect historical tweets, while Search API was limited by its free-to-use sandbox edition that imposes restrictions on collection limits and time spans based on queries used (Twitter, 2021a). In contrast, Academic Twitter offers greater flexibility in collecting historical data, with a collection limit of 10 million tweets per month, effectively reflecting no collection limit (Twitter, 2021b). Academic Twitter satisfied all the established criteria despite not meeting the "ease of access" criterion set in the previous chapter. Therefore, Academic Twitter was accessed via the application process and ultimately selected it as the preferred extraction tool.

## Appendix 12: Full Kruskal-Wallis H Results (Chapter Six)

<i>Strategy</i>	<i>Hashtag</i>	<i>Mention</i>	<i>Hyperlink</i>	<i>Emoji</i>	<i>Word Count</i>
<i>Organisation Information</i>	2916.36	2496.92	2248.71	2473.52	2892.00
<i>LOC (non-social media)</i>	1908.45	1957.60	2980.01	2282.97	2238.76
<i>LOC (social media)</i>	2469.03	2585.37	2840.88	2558.61	2360.12
<i>Non-Organisation Information</i>	3162.09	1864.55	2446.24	2462.89	2680.50
<i>Interactive</i>	2843.56	2226.83	2853.06	2584.32	3022.38
<i>Work Recruitment</i>	2078.53	2291.67	2867.04	2618.50	2721.18
<i>Observational</i>	2581.01	2281.06	2962.26	2464.54	2901.99
<i>Service Use</i>	2769.65	2226.90	2817.33	3196.52	2984.54
<i>Networking</i>	1821.00	3916.77	918.62	2439.74	1365.84
<i>Transactional</i>	2265.16	1930.78	2879.83	2474.69	3327.19
<b>Sig</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>
<b>Kruskal-Wallis H</b>	639.683	1138.020	2213.484	150.705	657.628

## Annexes

---

### ANNEX I

#### Welsh Government Evaluation of HateLab Platform Pilot

The Welsh Government's Inclusion and Cohesion Team is responsible for gathering and helping form responses to emerging community and societal tensions at the local, regional and national level in Wales. Its Community Cohesion Programme funds a network of 8 Community Cohesion Coordinators across regions in Wales. As part of their work programme, the Community Cohesion team monitors community tensions in their respective regions, working with the appropriate partners to mitigate these issues as they arise. The Cohesion team provide monthly monitoring reports to Welsh Government but will also send ad hoc updates if the matter is urgent or could potentially escalate. We partnered with HateLab as we have a need to monitor the online social space as part of our Community Cohesion team assessments. Social media, in particular Twitter, is a key space for the discussion of local and regional issues, and without HateLab's technology we had no programmatic way to reliably assess variation in online tensions that relate to offline events.

#### *Monitoring exercises using the HateLab Platform:*

As part of this pilot, we wanted to explore how the HateLab platform could potentially help supplement and enhance the work of the Cohesion teams. Since October 2021, whenever we have been alerted to emerging or potential community tensions that we felt appropriate, we have created queries on the platform to track these developing situations for any flashpoints which could result in hateful communications on social media.

Therefore, most of our activity during the pilot has been informed by intelligence supplied by our Cohesion teams. We feel these 'on the ground' sources lend themselves well to the platform, as many of the incidents tended to be localised, with

trackable and unique words useful for searching, such as Welsh place names or locations. This has helped to ensure in some cases that we have not collected as much irrelevant information. In some cases, the attached location was one which could be found in multiple countries, sometimes more than one in the same country, such as Newport. With the support from Dr. Arron Cullen, we looked at how to refine our searches.

Regarding the types of incidents, we have used the platform to follow, it is difficult to provide a case study as the incidents chosen were highly sensitive and in most cases the incidents are ongoing, whether unresolved or still under investigation. In general, we have used the platform to monitor very localised situations/incidents, which in every case, was not a hate incident, but had the potential to attract negative attention and hateful comments online. This approach has meant many of our queries did not identify hateful comments as the situations thankfully did not escalate to that point. This frequent outcome reflected our intended use of the platform.

For us, the value of the platform was not just in identifying hate speech, but also to monitor where tensions *did not* result in hate speech online. An example of this would be the migration of Ukrainians to the UK following the Russian invasion in February 2022. We used HateLab to monitor signs of any negative reaction to the news that Ukrainian families would be relocated to Wales, as there was a chance that far-right groups might capitalise on the situation with anti-migrant rhetoric. Concerns were increased following an incident in Hermon, Pembrokeshire where Ukrainian flags were torn down, as reported by the BBC. While we ran these queries, HateLab did not pick up any examples of hate speech online via the queries we set up. This tallied with the general feeling of goodwill towards Ukrainians that was supported by numerous other sources.

Likewise, there was an incident where a far-right group wrongly identified a hotel as a location where asylum seekers would be accommodated. The use of hotels to accommodate asylum seekers is an area which attracts interest from several far-right groups, who use these cases to cause division in local communities. We used HateLab to monitor whether this misinformation would spread and result in localised

tensions or hate speech. Again, it did not, which for us was a positive and provided additional reassurance that the situation had passed without escalation.

In those cases where situations did escalate, the platform showed us how quickly social media accounts from outside the community/area can take an interest in an incident or developing situation. For example, a protest about a particular cause will likely attract other like-minded accounts via hashtags or influencers. In these cases, we saw far more activity as more users were attracted to the developing situation.

As previously mentioned, we are aware that social media users have become more savvy in the way they direct abuse, and often do not use openly hateful language, instead choosing to use more coded words. An example would be the use of 'Ok groomer' towards the LGBTQ+ community and people that speak in favour of relationship and sexuality education in schools. The search function on HateLab provides a way of homing in on these terms through the lens of the anti-LGBTQ+ classifiers and provides us with a method of widening our searches and generally allowing us to be more dynamic to developing terms or words.

Going forward, we have considered how the platform could help us to respond with counter messaging/counter narratives. The word cloud function effectively pulls together common phrases, which in most cases are not overly hateful, but can support or promote stereotypical tropes and 'dog whistle' messaging. For example, our queries relating to asylum seekers have picked up the term 'economic migrants' which pushes a trope around people moving to the UK for financial reasons rather than them seeking sanctuary. With some co-ordination with partners, we would like to explore the use of targeted communications to counter such messaging.

## **ANNEX II**

### **Additional Evaluations of HateLab Platform Pilo**

#### **Galop (LGBTQIA+ anti-violence charity)**

For the last 35 years Galop has been working to support the rights of LGBT+ people in the UK, through our work on hate crime, domestic abuse, sexual violence and policing. Our professional casework and helpline services focus on empowerment-based support, advice and advocacy to people facing online and offline abuse and violence. We also deliver policy work addressing hate crime by producing research, guidance documents and training. Additionally, we work closely with government, criminal justice, charity and academic bodies to provide advice on issues relating to hate crime. Within the UK we lead the Community Alliance to Combat Hate, an intersectional partnership of leading anti-hate crime charities. Internationally we work closely with the European Commission, OSCE and overseas LGBT+ organisations to stand against online hate speech and hate crime.

*Monitoring exercises using HateLab platform:*

The HateLab platform was used by Galop for several months, supported by Dr. Arron Cullen. During this period, Galop ran several different monitoring exercises to track anti-LGBT hate online. These exercises were conducted using common and fewer common hashtags being used by those with anti-LGBT+ views. A sampling exercise was conducted using the platform to gather these hashtags, as well as using common hashtags of abuse from Galop's own service users. Once gathered the monitoring exercises were conducted. The exercises allowed Galop to pick up unknown hashtags that were being used online by those with anti-LGBT+ attitudes, which provided a steppingstone in discovering and tracking a wider set anti-LGBT+ beliefs than was previously possible without the platform.

For example, Galop ran a 'Monkeypox' monitoring exercise for several months using the platform. This exercise allowed us to identify when there were spikes in homophobic and transphobic comments being made online and to match these against offline world events. Using the platform, we saw a spike of homophobic comments against gay men when news articles on the outbreak were trending. This fed into Galop's narrative that there was a strong belief online that 'Monkeypox' was a 'gay' illness, fuelled by certain media outlets, that were helping generate homophobic stigma.

The HateLab platform proved to be a powerful research tool and Galop was very grateful to be part of the pilot. The learning from the pilot showed that the platform can provide strong evidence on online hate against our communities that is often left undetected. Overall, the HateLab platform is a brilliant resource.

*Billie Boyd, Advocacy and Support Manager, Galop*

## **National Online Hate Crime Hub**

The UK National Online Hate Crime Hub was established by the Home Secretary in 2017 to tackle online forms of hate crime, that increased dramatically in the aftermath of the 2016 referendum vote on the future of the UK in the EU. It acts as the point of contact for all victims of online hate crime, and produces intelligence reports.

### *Monitoring exercises using HateLab platform:*

The provision of the HateLab platform, co-created with the Hub, has fundamentally changed the way we monitor the spread of hate speech during national events. Prior to the platform, Hub staff relied on the Twitter platform interface to gather evidence on the ebb and flow of hate speech around events, such as the referendum vote on the future of the UK in the EU. This proved to be an inadequate method of generating the required insights to track and respond to the problem. During live operations, we were quickly inundated with irrelevant information and failed to capture hate speech in a systematic and reliable way. Through our close collaboration with HateLab, we have co-created technical solutions to overcome these problems. The platform employs sophisticated Machine Learning algorithms to automatically classify hate speech across recognised characteristics at scale and speed, and displays results via a range of visualisation tools (frequency chart, top hashtags, topic clusters, geo-location, networks etc). This ensures the Hub can monitor the production and spread of hate speech around events in a robust and reliable way.

To date, we have used the platform to monitor hate speech around key moments of the Brexit process, including the abuse of MPs, around terror attacks, and most recently around LGBTQ+ pride month. During these events the platform has allowed the staff in the Hub to better understand the dynamics of hate speech propagation, leading to improved response times, better support for victims and more effective allocation of resources.

Our latest monitoring exercise around LGBTQ+ pride month in the UK was conducted because our intelligence suggested we would see higher levels of online hate speech during this event than we would typically expect to see throughout the rest of the year. The HateLab platform allowed us to clearly track online hate at the regional and national level during the event. We were able to determine trends on an hourly basis, generate reports centred around keywords and trending hashtags used by those writing the messages, allowing us to create a fuller view of the issue. The platform was an invaluable resource providing a swift analysis of hateful and harmful material in popular discourse allowing us to assess and respond to community threats in real-time.

Overall, the platform provides professionals with the best informed assessment of societal tensions and, combined with nearly a decade of observations of the analytics of social media societal tensions, community intelligence and crime trend data, it enables police managers and partners to make the best-informed decisions on deployment and preventative interventions.

*Paul Giannasi, National Policing Lead for Hate Crime*

### **EE, BT and Saatch & Saatchi**

EE, BT and Saatchi & Saatchi required data on hateful online messages sent to professional footballers for national anti-hate campaigns (Hope United) during the Women's Euros and the World Cup. Marketing assets using data the HateLab Platform included apparel, TV, OLV, OOH, and Digital.

“HateLab were an essential partner for our Hope United campaigns around the Women's Euros and the World Cup. Their knowledge of online harms and their



award-winning AI was of huge value and gave us depth of understanding and credible insights we couldn't have otherwise achieved. BT and EE would not hesitate work with HateLab on future projects.”

*Alice Tandler, Group Head of Brand Strategy BT Group*

“HateLab were instrumental in providing the data foundation for our Women’s Euros campaign for EE Hope United. We were introduced to them through the data driven shirt design project, working with The Mill Experience; whereby players’ individual data histories of online abuse impacted the unique designs of their Hope United kits. The team’s tracking of hate across the tournament allowed us to confidently talk about the levels of misogynistic hate our Hope United players received during the Euros and reflecting that in reactive press, digital out of home and social which ran over the finals weekend. We are working with the team again, on the World Cup, and the data tracked during that period is a creative jumping off point for a current project – orientated around tackling homophobia within the men’s game.”

*Sophie deGraft-Johnson, Business Leader One EE, Saatchi & Saatchi*