A Discourse-Analytic Approach to the Study of Information Disorders:

How Online Communities Legitimate Social Bonds When Communing Around Misinformation and Disinformation

Olivia Inwood

A thesis in fulfilment of the requirements for the degree of Doctor of Philosophy

School of the Arts and Media

Faculty of Arts, Design and Architecture

February 2023

ORIGINALITY STATEMENT

☑ I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

COPYRIGHT STATEMENT

AUTHENTICITY STATEMENT

☑ I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis.

UNSW is supportive of candidates publishing their research results during their candidature as detailed in the UNSW Thesis Examination Procedure.

Publications can be used in the candidate's thesis in lieu of a Chapter provided:

- The candidate contributed **greater than 50%** of the content in the publication and are the "primary author", i.e. they were responsible primarily for the planning, execution and preparation of the work for publication.
- The candidate has obtained approval to include the publication in their thesis in lieu of a Chapter from their Supervisor and Postgraduate Coordinator.
- The publication is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in the thesis.

☑ The candidate has declared that **some of the work described in their thesis has been published and has been documented in the relevant Chapters with acknowledgement**.

A short statement on where this work appears in the thesis and how this work is acknowledged within chapter/s:

> Several publications have emerged based on research undertaken for this thesis. Inwood and Zappavigna (2021a, published in Discourse & Communication) emerged from the analysis of the Momo Challenge case study comments from Chapter 7. It introduces the method for identifying textual personae in YouTube comments via an appraisal and affiliation analysis. Additionally, it introduces the bond cluster diagram as a way of mapping out the key social bonds shared by particular textual personae and how these bonds relate to other textual personae. Inwood and Zappavigna (2023, published in Social Media + Society) emerged from the analysis of the Notre Dame Fire case study comments from Chapter 7. It also introduces the method for identifying textual personae in YouTube comments via an appraisal and affiliation analysis, and additionally explores the key legitimation strategies adopted by different textual personae. Inwood and Zappavigna (forthcoming, Routledge Handbook of Discourse and Disinformation) and Inwood and Zappavigna (2022a) are other publications that focus on explaining the process for identifying textual personae in YouTube comments and emerged from the method developed in Chapter 4.
>
> Inwood and Zappavigna (2022b, published in The Communication Review) emerged from the analysis conducted in Chapter 6 regarding the Notre Dame fire video transcripts and the analysis conducted on the Notre Dame fire replies to comments in Chapter 7. This work focused on explaining how an appraisal and affiliation analysis can illustrate how patterns of evaluation are expressed in language. As well as how these function in terms of aligning ambient audiences with particular values, to offer an additional perspective on issues of information disorder that does not attempt to homogenise the multiple reasons why people engage in such hateful behaviour. A book chapter currently under review “The legitimation and delegitimation of values in white supremacist and conspiratorial discourse” also emerges from the analysis undertaken in Chapter 6 in relation to the legitimation analysis conducted on the Notre Dame fire video transcripts. Lastly, a journal article under review “The legitimation of screenshots as visual evidence in social media: YouTube videos spreading misinformation and disinformation” emerged from the affiliation and legitimation analysis conducted in Chapter 6 and outlined in Chapter 4 regarding the visual content of the YouTube videos from both the Momo Challenge and Notre Dame Fire case studies.
>
> All these publications have been acknowledged throughout the thesis in the relevant chapters they are drawn from. The published articles and book chapters are the direct outcome of research related to Olivia Inwood’s PhD thesis, supervised by Associate Professor Michele Zappavigna. Olivia Inwood is the lead author of all the articles and book chapters included and Associate Professor Michele Zappavigna is the co-author. Olivia Inwood collected and analysed all the data for this research and wrote the majority of the research article. Olivia Inwood and Associate Professor Michele Zappavigna have both signed the UNSW authorship agreement in order to confirm this statement.

Candidate's Declaration

I declare that I have complied with the Thesis Examination Procedure.

ABSTRACT

Information disorders have become prevalent concerns in current social media research. This thesis is focused on the interpersonal dimension of information disorders, in other words, how we can trace, through linguistic and multimodal analysis, the social bonding that occurs when online communities commune around misinformation and disinformation, and how these social bonds are legitimated to enhance perceived credibility. Social bonding in this thesis refers to a social semiotic perspective on the shared values that communities use to construe alignment with others. False information can spread when groups have a shared vested interest, and so information disorders need to be elucidated through an investigation of sociality and bonding, rather than via logical points alone. The term 'information disorder' encompasses the spectrum of false information ranging from misinformation (misleading content) to disinformation (deliberately false content), and it is within this landscape of information disorders that this thesis emerges. Two key forms of social semiotic discourse analysis were applied to a dataset of YouTube videos (n=30) and comments (n=1500): affiliation (analysis of social bonding) and legitimation (analysis of resources used to construct legitimacy). The dataset constituted two contrasting case studies. The first was non-politically motivated misinformation in the form of an internet hoax leveraging moral panic about children using technologies. The second was politically motivated conspiracy theories relating to the Notre Dame Cathedral fire. The key findings of this thesis include the multimodal congruence of affiliation and legitimation across YouTube videos, the emergence of technological authority as a key legitimation strategy in online discourse, and the notion of textual personae investigating the complex array of identities that engage with information disorders in comment threads. Additionally, six macro-categories were identified regarding communicative strategies derived from comment threads: scepticism, criticism, education and expertise, nationalism, hate speech, and storytelling and conspiracy. This shows not only how information disorders are spread, but also how they can be countered. The method outlined in this thesis can be applied to future interdisciplinary analyses of political propaganda and current global concerns to develop linguistic and multimodal profiles of various communities engaging with information disorders.

DEDICATION

I dedicate this thesis in loving memory to my grandparents: Clare Inwood, Beau Inwood, Emilija Marcinkienė and Steponas Marcinkus

ACKNOWLEDGEMENTS

I pay my respects to all the Traditional Owners of the lands in which I have lived as a settler and second-generation migrant. I acknowledge the Bedegal people who are the Traditional Owners of the unceded territory of the UNSW campus and the Wiradjuri people who are the Traditional Owners of the unceded territory of Central Western NSW. Aboriginal and Torres Strait Islander peoples are also the first knowledge holders of this continent, whose cultures are among the oldest living in human history.

I have only been able to complete this thesis due to the number of privileges I've experienced and amazing people I have met in my life.

Firstly, my heartfelt gratitude goes to my supervisor Associate Professor Michele Zappavigna. As a young honours student in 2018 I remember I was nervous about having a meeting with an academic, but from the very start Michele was very kind, patient, and generous in helping me learn and develop confidence. These wonderful experiences during my honours year encouraged me to pursue a PhD. I am so grateful for Michele continuing to supervise me during my PhD candidature and for helping me secure funding. Michele has given me many opportunities during my PhD candidature, such as co-authoring publications and teaching opportunities. She has provided detailed feedback on all my work and given so much advice, both academic and personal, in times of need. Thank you, Michele, you are an inspiration.

Many amazing communities have provided support to me over the years that I would now like to take the time to thank. Thank you to the Australian Systemic Functional Linguistics Association (ASFLA), a community that has supported me throughout my PhD candidature. Thank you to Dr Peter White my secondary supervisor and Professor Louise Ravelli my PhD reader, for their incredibly helpful feedback on my work. I also had the opportunity to attend ASFLA seminars and events, and I would like to thank Professor Jim Martin, Associate Professor Sue Hood, Dr Yaegan Doran, Professor Theo van Leeuwen, Associate Professor Shooshi Dreyfus and Associate Professor Helen Caple for all their amazing contributions to ASFLA. A massive thank you as well to all the wonderful friends I made through the ASFLA community, including, Nataliia Laba, Dr Awni Etaywe, Dr Giselle Newton, Dr Gaga Stosic, Dr Xiaoqin Wu, Nida Tahseen, Gyeyoung Lee, Dr Lorenzo Logi, Georgia Carr, Annie Hellwig and Jiani Chen. Thank you as well as Dr Eszter Szenes and Dr Mohammad Makki for their interest in my work.

Thank you to the Commonwealth of Australia and the Australian people for funding my research. Thank you to the Defence Science and Technology Group for funding my research and for their continual collaboration. It is an honour that my research can have a positive impact.

Thank you to Forbes High School, I will never forget the immense support my teachers gave me, I wouldn't have made it to university without them. In particular, I would like to thank Anne Glynn, Jane Hill, Sue Cunningham, Jacqueline Flannery, Anna Townend, Trudi Weyman and Fiona Isaacs for all their care and for continuing to follow my life updates.

I thank the community of UNSW, I was only able to afford to study in Sydney as an 18 year old due to the generosity of donors contributing to equity scholarships for rural students. Thank you to Fig Tree Hall, the college I lived and worked at for nine years. At times research and writing can be a lonely pursuit and I am so grateful I had a home at Fig Tree Hall. I have met the most wonderful residents and colleagues, and have lifelong memories I will cherish. In particular, I would like to thank Rasheed Khan and Isabelle Creagh for all their support and for believing in me enough to hire me. Thank you, Poppy, for your amazing positive energy. And thank you Jenny Power for being the ultimate UNSW Colleges legend.

Thank you to the Faculty of Arts, Design and Architecture for providing financial support and research assistant opportunities. Thank you to the wonderful academics that taught me during my undergrad, including, Associate Professor Sean Pryor, Professor Ramaswami Harindranath and Associate Professor Michael Richardson. Thank you as well to the students I taught in ARTS2091 and ARTS2093, who kept my enthusiasm of learning alive and provided generous feedback.

Last but not least, I thank my parents for everything they have given me. My parents have experienced a tough life but have always cared for me and treated everyone with kindness. My Mum left her entire life in Lithuania to move to Australia and start anew. As a second-generation migrant I benefit from my Mum's sacrifices. My Dad as a farmer has battled through a decades-long drought, and then floods, including one that damaged my childhood home. My parents inspire me to do work that I personally find meaningful, work hard, and not take anything for granted.

TABLE OF CONTENTS

VOLUME II: APPENDICES

LIST OF TABLES

LIST OF FIGURES

1. CHAPTER ONE INTRODUCTION: INFORMATION DISORDERS AND THE LEGITIMATION OF SOCIAL BONDS ONLINE

1.1. Information Disorders and Social Bonding on YouTube

This thesis is interested in the interpersonal dimension of information disorders, in other words, how we can trace, through linguistic and multimodal analysis, the social bonding that occurs when online communities commune around misinformation and disinformation, and how these social bonds are legitimated to enhance perceived credibility. Social bonding refers to the process through which communities enact alignments through the shared values construed in their language. This thesis adopts a social semiotic perspective on this bonding in order to considers how values relating to false information spread when groups have a shared vested interest (e.g. a hatred of immigrants or a moral panic about a new technology) (Carlson, 2020; Murphy, 2022; Smith et al., 2020).

This thesis is motivated by recent research suggesting that we need to combat information disorders through investigating sociality and bonding, as they cannot be combatted via logical points alone (Abraham, 2014; Van der Linden and Roozenbeek, 2020; Vraga et al., 2019). It thus proposes a method to investigate the interpersonal dimension of information disorders via a social semiotic discourse analysis applied to a dataset of YouTube videos and comments. This method applies both an affiliation analysis (of social bonds) and legitimation analysis (of resources used to construct legitimacy) to understand the interpersonal core of information disorders. Social media networks like YouTube are visually-orientated and so this thesis also seeks to explore the visual modality alongside language, and contribute to the less researched area of visual information disorders (Heley et al., 2022; Hemsley and Snyder, 2018).

The spread of deliberate disinformation, and the ease with which anyone can cause confusion among audiences during breaking news events, have become prevalent concerns in current social media research. Concepts such as deception (Chisholm and Feehan, 1977), lying (Barnes, 1994; Bok, 1978; Ekman, 1985), and propaganda (Biddle, 1931) have been the ongoing focus of research throughout the 20th century. However, since the 2016 US Presidential Election, the term 'fake news' has gained extensive prominence in mainstream discourse, and social media platforms have been increasingly criticised for their handling and manipulation of information (Allcott and Gentzkow, 2017). Prior to 2016, 'fake news' was largely used in academic work to refer to satirical content (Reilly, 2012), however the term has since increased in scope, and some researchers have incorporated manipulation, falsification, and propaganda into its definition (Tandoc Jr et al., 2018). During the 2016 Presidential Election former President Trump regularly used the term 'fake news' to criticise any unfavourable media coverage or viewpoints, hence this term became an interest of scholarly inquiry in light of this socio-political climate (Holan, 2017). Farkas and Schou (2018) highlight the convoluted nature of 'fake news' when describing how it has been articulated in three different ways. This includes as a critique of digital capitalism, a critique of right-wing politics and media, or a critique of the liberal and mainstream media (Farkas and Schou, 2018: 300).

Due to the imprecision of these definitions, Wardle and Derakhshan (2017: 5) propose a more comprehensive framework which they term "Information Disorder". This framework encompasses: misinformation (misleading content), disinformation (deliberately false content), and malinformation (content with an intent to harm such as information leaks and hate speech) (Wardle and Derakhshan, 2017). These categories allow the complexity of deceptive and false information practices to be explained more precisely. The framework also establishes the umbrella term 'information disorder' to encompass the range of false information practices one can encounter. This is proposed as a more neutral term than the politically charged 'fake news'. Avoiding terms for describing false information that have a different meaning depending on the discipline or context used is particularly important when considering how difficult it can be to

distinguish misinformation from disinformation on social media platforms (Søe, 2018). This thesis acknowledges the complexity of discussing information disorders, and therefore the complexity in researching such a phenomenon.

It is within this landscape of information disorders that this thesis emerges. The thesis takes an exploratory qualitative approach to investigating the interpersonal dimension of information disorders. This is explored through a detailed analysis of YouTube videos and comments from two case studies representing politically-motivated and non-politically-motivated information disorders. YouTube was selected as the data source since it features an abundance of multimodal social media content, allowing both video content and comments to be analysed. YouTube features a participatory framework (Dynel, 2014) enabling the interactions of social media users in comment threads to be analysed in multiple ways. For social media users this means that YouTube comments can serve as a space to pseudonymously provide commentary on an issue addressed in a video but avoid directly engaging with other user comments. Conversely, YouTube comments also serve as a space to directly reply to and name users in interactions that extend beyond the content of the video. For researchers, YouTube provides a wealth of freely available public data that stretches across "international and inter-generational audiences" (Thelwall, 2018: 304). While internet research ethics are a current area of extensive and on-going debate, this thesis aligns with the approach of a critical-realist internet research ethics. This means that public data can be used for research if the researcher intends to use this data for public good, does not manipulate the data of individuals for profit, and makes all reasonable attempts to anonymise the data of private individuals (Fuchs, 2018). YouTube data provides discourse analysts with the ability to closely analyse data from specific historical moments and can therefore provide public good in terms of research value. Additionally, YouTube is a less studied social media platform in regards to issues of misinformation and hate speech (Lewis, 2018; Matamoros-Fernández and Farkas, 2021), and so there is potential to contribute to further qualitative analyses of YouTube in this regard. Therefore, this thesis will contribute to the research of the interpersonal dimension of information disorders on YouTube via a linguistic and multimodal discourse analysis of social bonding.

1.2. A Discourse-Analytic Approach and Interpersonal Meaning

A discourse-analytic and social semiotic perspective on interpersonal meaning is centred on how relationships are enacted through language and other semiotic resources (Halliday, 1978). This includes considering how authorial identities in texts align or dis-align with others and how communities construct shared feelings and values (Martin and White, 2005). The broad focus of this thesis is on the interpersonal dimension of information disorders, in other words, how these authorial identities and communities' bond around shared values in order to propagate their construal of information disorders. This section will provide a brief orientation to the theoretical and analytical methods informing this research that are explained in more detail in Chapter 3.

This thesis adopts the specific discourse-analytic approaches of Social Semiotics and Systemic Functional Linguistics (hereafter SFL). Social semiotics is the study of how semiotic resources are construed in "specific historical, cultural and institutional contexts" (Van Leeuwen, 2005: 3). According to this functional resource-oriented perspective language is not a set of prescribed rules but rather "a resource for making meaning" (Halliday, 1978: 192). Semiotic resources include modalities other than language such as visual images and encompasses any of "the actions and artefacts we use to communicate" (Van Leeuwen, 2005: 3). In this sense, the focus of this thesis on the visual as well, aligns with a social semiotic multimodal discourse perspective.

SFL, as social semiotic theory, is a linguistic approach that studies meaning-making resources in language according to their social context (Halliday, 1978: 122). It is defined by its systematic and detailed methods for studying language patterns, and offers methods for analysing language across the different strata of language from phonology to lexicogrammar and discourse semantics. A core concern of SFL as an 'appliable linguistics' (Halliday, 2008) is applying this analysis to practical situations in order to address social issues. Thus, the aims of SFL align with many core concerns in communication and discourse studies, that is, the focus on studying social contexts and applying rigorous theoretical models to understanding social issues.

SFL is a multi-perspectival model, that considers language as a resource for making ideational, interpersonal, and textual meanings. The ideational is concerned with construing experience, the interpersonal is concerned with how relationships are enacted, and the textual is concerned with "how ideational and interpersonal meanings are distributed in waves of semiosis, including interconnections among waves and between language and attendant modalities" (Martin and White, 2005: 7). Altogether, these three perspectives on the broad functions of language are known as metafunctions (and will be explored further in the Chapter 3). As previously stated, the focus of this thesis is on the interpersonal dimension of information disorders. In SFL terms, the interpersonal is concerned with the negotiation of social relations. Nonetheless, these metafunctions all work together in order to construe meaning in language.

This thesis uses the SFL framework of affiliation (Zappavigna, 2018), to identify how social bonding occurs in language, and considers how the fusion of ideational and interpersonal meanings form the discursive basis of social bonds. The multimodal discourse framework of legitimation (Van Leeuwen, 2007) is also applied to understand how these social bonds are legitimated in discourse through choices in language and other resources. These frameworks can be applied to both contexts involving direct interaction between people and to instances where a text receives no reply. Thus, these frameworks are applied in the case studies both to the communication in the YouTube videos, as well to the comment feeds that occur with these videos. A YouTuber's vlog can aim to propagate social bonds that audiences can then rally around, or dialogic communication can occur as in how two users might reply to each other's comments on a YouTube video. In this sense, the enactment of social bonds can occur across various domains and in monologic and dialogic formats.

It is important to reflect on what it means to analyse information disorders from a social semiotic perspective in order to clarify what a social semiotic analysis can uncover and its consequent limitations. The analysis undertaken in this thesis is concerned with what can be revealed by the meanings people make with semiotic resources in YouTube videos and comments. In accord with social semiotic theory, it is thus not concerned

with individual psychological states or intentions of the YouTubers in the production of this content. The case studies were selected on the basis of their position in the broader social context as politically or non-politically motivated, and as *a priori* constituting untrue information. Rather than identifying *what* is false, this thesis offers a means to understand *how* people bond around false information, and offer insights into *why* this occurs by identifying specific social bonds.

1.3. Research Questions

The overall aim of this thesis is to enhance understanding of the interpersonal dimension of information disorders via linguistic and multimodal discourse analysis methods. More specifically, these aims are to:

1. Understand the key affiliation and legitimation strategies used in the language and visual communication of YouTubers engaging with information disorders.
2. Identify and analyse the key social bonds that textual personae share or contest in comments when they respond to YouTube videos that engage with information disorders.
3. Compare and contrast textual personae in politically motivated information disorders (e.g. politically-triggered conspiracy theories) versus non-politically motivated information disorders (e.g. internet hoaxes).

These aims are addressed using two forms of discourse analysis: affiliation (the study of social bonding) and legitimation (the study of legitimating features in language and other modalities). These methods can aid in the exploration of the interpersonal dimension of information disorders, by investigating how social bonds are legitimated to enhance perceived credibility. The identification of particular patterns of affiliation strategies and social bonds (i.e. the shared value that is propagated), arises in the formation of textual personae, as will be further explained in Chapter 4. This concept of textual personae allows further discussion regarding the meaning of particular social bonds in the context of information disorders, in other words, how certain groups position themselves in relation to discourse about information disorders and the

linguistic strategies they adopt to achieve this, as well as how these groups convince each other that their information is correct. By comparing textual personae across both case studies, representing different macro-genres of information disorders, this then allows an exploration of how textual personae work in different contexts.

Based on these aims, four specific research questions evolved:

1. How can the interpersonal dimension of information disorders spread on YouTube be illuminated via a linguistic and multimodal analysis of patterns of affiliation and legitimation?
2. How are social bonds legitimated in the transcripts and visual content of YouTube videos that spread information disorders?
3. What different textual personae emerge in the comment sections of YouTube videos that spread information disorders? Can these textual personae be grouped according to social bonds and legitimation strategies?
4. What similarities and differences emerge when comparing textual personae in non-politically motivated information disorders versus politically motivated information disorders?

The first research question addresses the broad overarching aims of this thesis; to explore the interpersonal dimension of information disorders on YouTube via the linguistic and multimodal analysis of patterns of affiliation and legitimation. The second research question relates more specifically to the multimodal affiliation and legitimation analysis of YouTube videos in Chapters 5 and 6. The third research question relates specifically to Chapter 7 regarding the exploration of textual personae in YouTube comment sections. The final research question is another overarching research question, that reflects on the significance of this thesis. This question ties together the two case studies analysed to explore the similarities and differences among textual personae in non-politically-motivated information disorders versus politically-motivated information disorders.

Overall, these research questions are designed to explore the interpersonal dimensions (social bonding and the legitimation of social bonding) of information disorders on YouTube via linguistic and multimodal discourse analysis frameworks. This exploration allows an understanding of how online communities commune around information disorders, and the development of textual personae in order to enhance this understanding. The research questions align with particular chapters in this thesis as will be explored in Section 1.5. Additionally, the order of the research questions reflects the progression of the research across the thesis.

1.4. Overview of Case Studies: Non-Political versus Political Discourse

This thesis is centred around two case studies: The Momo Challenge (hereon, the Momo Challenge case study) representing non-politically motivated information disorders and The Notre Dame Cathedral Fire (hereon, the Notre Dame Fire case study) representing politically motivated information disorders. These case studies were selected on the basis that the YouTube videos could be defined as instances of viral or spreadable media (Jenkins et al., 2013) and involved information that was patently false. They received significant mainstream media attention, YouTube video views, and were frequently searched for on Google. Additionally, the videos featured clearly identifiable instances of information disorder, albeit different aspects of information disorder, meeting the criteria of containing information that was verifiably untrue. It was important for the falsity of the videos to be predefined, rather than established in the data analysis phase, in order to have enough information disorder data to study, and since our focus is on social bonding rather than on detecting false information. These case studies will now be elaborated upon in more detail.

The Momo Challenge refers to an internet hoax about children receiving threatening messages on WhatsApp and being shown images of a frightening figure called Momo (as illustrated in Figure 1-1). Other variations of the story also reported that footage of the frightening figure was spliced into YouTube Kids videos such as 'Peppa Pig' videos. The Momo Challenge was incorrectly linked to the suicide of children from various parts of the world. No evidence has been discovered or statements from law enforcement

agencies that directly links the Momo Challenge to the death of any children[1]. The Momo Challenge has been discredited by the majority of major media outlets and YouTube has stated that it did not discover any Momo Challenge spliced videos. Thus, accounts of the Momo Challenge most commonly refer to Momo as an 'internet hoax' or 'internet urban legend'. The Momo Challenge is unique in that it is one of a few internet hoaxes that have achieved a viral status across social media and broadcast media, and it also represents continued concerns regarding online safety and moral panics due to a lack of knowledge about internet culture and younger generations. The Momo Challenge is non-political because it does not consider discourses surrounding the politics of certain countries, rather it focuses on the nuances of the internet cultures of different countries.

Figure 1-1 – Image of Momo (creative commons licence)

The Notre Dame Fire case study involved YouTube videos about a fire that occurred on 15th April 2019 which caused extensive damage to the Notre Dame Cathedral's roof and upper walls, and led to the collapse of the spire. French prosecutors declared that there was no evidence of the fire being a deliberate act and that it most likely was caused by an electrical fault[2]. The news event received a global reaction due to the historical and cultural significance of the cathedral, and was one of the most googled news events of 2019 worldwide, with images such as the one shown in Figure 1-2 becoming iconic imagery during this time. Due to this global impact, conspiracy theories about how the

[1] Further details and references about the Momo Challenge are provided in Chapter 4 and Chapter 5.

[2] Further details and references about the Notre Dame Fire are provided in Chapter 4 and Chapter 6.

fire started proliferated on social media, as well as white supremacist hate speech. For example, posts emerged that blamed Muslims for the fire, labelling it as a terrorist act. It is this global attention that the Notre Dame Fire attracted and its political significance (i.e., tensions regarding multiculturalism, immigration, and criticisms of the French government) that makes it worthy to consider as a case study for this thesis. It is a highly political case study because of these tensions, and its reflection on French politics.

Figure 1-2 – Image of Notre Dame Cathedral Fire (creative commons license)

Overall, these case studies represent events of historical importance because they received significant global attention online. Visually, these two case studies are also powerful, because of their iconic visual imagery: the image of the Notre Dame Cathedral on fire, and the image of the frightening figure of Momo. It could be argued that without these powerful visual images, these case studies would not have had such a viral reaction. The selection of these two case studies allows the spectrum of information disorders to be explored, ranging from internet hoaxes, conspiracies, and white supremacist discourse.

1.5. Thesis Structure

The structure of this thesis is designed to introduce the key dimensions of information disorders and to explain how an affiliation and legitimation analysis can serve to illuminate the social bonding at the heart of these disorders. It then undertakes an analysis and comparison of two case studies of YouTube information disorders, before

discussing the implications and main findings of the research. Below is a brief synopsis of each chapter in this thesis.

Chapter 2: Literature Review – Interpersonal Perspectives on Information Disorders – The literature review critically evaluates studies in information disorders. It begins with an exploration of the predecessors to studies in information disorders online, that is, models about deceptive communication, propaganda, and lying. The information disorder framework is then explored, and an overview of the extent to which misinformation, disinformation, and malinformation have been analysed across multiple disciplines from an interpersonal perspective is undertaken. The chapter concludes with a reflection on how this thesis can make a contribution to the study of information disorders via its focus on the legitimation of social bonds.

Chapter 3: Theoretical Framework – A Discourse-Analytic Approach Grounded in Social Semiotics – The theoretical framework chapter locates the specific disciplinary orientation of the thesis. It provides an overview of social semiotics, the main theoretical framework of this thesis, as well as a brief history of Systemic Functional Linguistics (SFL). Key SFL concepts that the thesis draws upon are explained: the appraisal and affiliation frameworks, discourse iconography, textual personae, instantiation, individuation, and genre. Multimodal discourse analysis is also detailed, alongside an explanation of key concepts in legitimation, identity, multimodal computational studies, and approaches to analysing social values. The chapter ends with a reflection on how these theoretical frameworks inform the research questions of this thesis.

Chapter 4: Methodology – Analysing YouTube Videos and Comments – The methodology chapter will explain in detail how the data was collected, sampled, and analysed. Firstly, the methodological motivations and research questions are explored, followed by an explanation of how the data was collected and the kinds of case studies analysed. The procedure for undertaking the affiliation and legitimation analyses is detailed. The chapter concludes with a process diagram summarising the entire data collection and analysis process.

Chapter 5: The Momo Challenge - Non-Politically Motivated Moral Panics and Internet Hoaxes – This chapter details the analysis of the YouTube videos undertaken in the first case study on the Momo Challenge, focusing on the discourse analysis of the videos. It addresses the second research question regarding how social bonds are legitimated in the transcripts and visual content of YouTube videos that spread information disorders. The analysis begins with identifying the different macro-genres of the videos in the dataset. The multimodal affiliation and legitimation analysis of the videos is then presented according to the key targets of the legitimation and affiliation strategies. The significance of these findings will then be reflected upon.

Chapter 6: The Notre Dame Fire – Politically Motivated Conspiracies and White Supremacist Discourse – This chapter details the analysis of the YouTube videos in the second case study about The Notre Dame Fire. This chapter is structured in a similar manner to the previous case study and addresses the second research question of the thesis. It begins with analysis of the different macro-genres in the Notre Dame fire dataset and then presenting the multimodal affiliation and legitimation analysis of the videos according to the key targets of the legitimation and affiliation strategies, before reflecting on their significance.

Chapter 7: Comparing Case Studies – Personae in Political versus Non-Political Discourse – This chapter compares the results of the YouTube video comment analysis from the two case studies and addresses the third and fourth research questions. It will group textual personae according to their social bonds and legitimation strategies. The textual personae identified in the Momo Challenge case study are firstly explored, as well as a dialogic affiliation analysis on replies to initiating comments. This is followed by an exploration of the textual personae in the Notre Dame Fire case study, as well as a dialogic affiliation analysis on the replies to initiating comments. After identifying these personae via bond cluster diagrams and analysing their most common legitimation strategies, the personae across the two case studies are compared. The results are discussed in detail as well as the potential application of textual personae

analysis in other domains to enhance our understanding of information disorders online.

Chapter 8: Conclusion – A Multi-Layered, Interpersonal Approach to Studying Information Disorders – The conclusion chapter unifies the key results from across the thesis, as well as discussing the implications, limitations, and future directions of the research. The key findings are grouped according to four key themes that highlight the main contributions this thesis has made: the significance of conducting a multimodal affiliation and legitimation analysis on YouTube videos, the importance of technological authority across both case studies, the exploration of textual personae, and the comparisons between non-political versus political discourse in the context of information disorders. The implications of the thesis are then explored according to their relevance to three main fields of study: SFL and social semiotics, communication studies, and interdisciplinary research. The limitations of the study are discussed, before focusing on future directions of this research and the contributions this thesis can make to the study of information disorders from an interpersonal perspective.

1.6. Publications Arising from this Thesis

Several publications have emerged based on research undertaken for this thesis. Inwood and Zappavigna (2021) emerged from the analysis of the Momo Challenge case study comments from Chapter 7. It introduces the method for identifying textual personae in YouTube comments via an appraisal and affiliation analysis. Additionally, it introduces the bond cluster diagram as a way of mapping out the key social bonds shared by particular textual personae and how these bonds relate to other textual personae. Inwood and Zappavigna (2023) emerged from the analysis of the Notre Dame Fire case study comments from Chapter 7. It also introduces the method for identifying textual personae in YouTube comments via an appraisal and affiliation analysis, and additionally explores the key legitimation strategies adopted by different textual personae. Inwood and Zappavigna (forthcoming) and Inwood and Zappavigna (2022a) are other publications that focus on explaining the process for identifying textual

personae in YouTube comments with different case studies to the ones explored in this thesis, and emerged from the method developed in Chapter 4.

Inwood and Zappavigna (2022b) emerged from the analysis conducted in Chapter 6 regarding the Notre Dame fire video transcripts and the analysis conducted on the Notre Dame fire replies to comments in Chapter 7. This work focused on explaining how an appraisal and affiliation analysis can illustrate how patterns of evaluation are expressed in language. As well as how these function in terms of aligning ambient audiences with particular values, to offer an additional perspective on issues of information disorder that does not attempt to homogenise the multiple reasons why people engage in such hateful behaviour. A book chapter currently under review "The legitimation and delegitimation of values in white supremacist and conspiratorial discourse" also emerged from the analysis undertaken in Chapter 6 in relation to the legitimation analysis conducted on the Notre Dame fire video transcripts. Lastly, a journal article under review "The legitimation of screenshots as visual evidence in social media: YouTube videos spreading misinformation and disinformation" emerged from the affiliation and legitimation analysis conducted in Chapter 6 and outlined in Chapter 4 regarding the visual content of the YouTube videos from both the Momo Challenge and Notre Dame Fire case studies.

1.7. Contributions of this Thesis

Overall, this thesis is exploratory in nature and contributes to the construction of a methodological framework for understanding information disorders from an interpersonal perspective, drawing on tools in SFL and multimodal discourse analysis. It does not attempt to provide quantitative solutions to information disorders, nonetheless, there is potential for this method to contribute to interdisciplinary work into information disorders as will be further discussed throughout. Analysing how social bonds are legitimated provides valuable insights into how information disorders are spread and countered by various online communities. In addition, this research shows how the legitimation of social bonds increases the impact of these shared values via legitimation strategies in linguistic and multimodal forms.

2. CHAPTER TWO LITERATURE REVIEW: INTERPERSONAL PERSPECTIVES ON INFORMATION DISORDERS

2.1. Introduction

In order to understand the academic literature on information disorders, it is essential to consider the vast array of categories and definitions of problematic or false information that have developed over time in this research area. This chapter will begin by discussing key historical models regarding the study of information disorders that are relevant to the interpersonal perspective developed in this thesis (although not necessarily themselves using this term). It will consider how these definitions connect to the broader literature on information disorders. The chapter will then consider the prominence of research into information disorders from the 2016 US Presidential Election onwards. By unpacking the key components of the information disorder framework (Wardle and Derakhshan, 2017), namely misinformation, disinformation, and malinformation, examples will be provided of research analysing these components of information disorders, with a focus on explaining how sub-types of information disorders have emerged based on foregrounding interpersonal relations, in other words, how people and institutions position themselves in relation to the spread of false information. The second part of this chapter will provide examples of work on information disorders using algorithmic, linguistic, and multimodal methods. YouTube as a social network known for propagating information disorders will also be explored (Lewis, 2018; Matamoros-Fernández, 2017). Based on these foundations, the chapter will then demonstrate how this thesis will fill a current gap in knowledge concerning the interpersonal significance of how individuals and groups position themselves in relation to discourse surrounding information disorders on YouTube via linguistic and multimodal discourse strategies.

As briefly discussed in the *Introduction Chapter*, the interpersonal metafunction in SFL is concerned with the negotiation of social relations in discourse. The focus of this thesis is on how social bonds are negotiated and legitimated in the discourses of online

communities engaging with information disorders. The academic literature on information disorders is especially vast, encompassing multiple disciplines of study, and so this chapter is focused upon studies that examine the interpersonal dimension of information disorders, in other words, how people engage with each other in the scope of information disorders. Studies that are focused on identifying what is false without considering the interpersonal relations among people discussing information disorders are not the focus of this chapter. Rather, this chapter will primarily examine research about the interpersonal dimension of information disorders on social media.

2.2. Deceptive Communication, Propaganda, and Lying

In order to firstly encapsulate the study of information disorders, preceding studies in information disorders need to be examined. This consists of studies in deceptive communication, propaganda, and lying. These particular concepts were the focus of much academic research prior to the post-2016 interest in "fake news". Various models have been applied to understanding deceiving and lying across the fields of psychology, philosophy, and sociology. The most common model across the aforementioned disciplines for mapping out the distinctions between deceiving and lying, involves defining 'deception' as the 'superordinate' category, and 'lying' as the 'subordinate' category (Barnes, 1994; Bok, 1978; Castelfranchi and Poggi, 1994; Chisholm and Feehan, 1977; Frank and Svetieva, 2013; Kalbfleisch and Docan-Morgan, 2019). Deception consists of many 'acts'. It is a social interaction and dependent on the intentions of the deceiver to deceive someone. In contrast, lying consists of one act, must be 'stated', and does not depend on the deceiver's success or intention.

Gupta et al. (2012) and McGlone and Knapp (2019) have developed taxonomies for classifying deception. McGlone and Knapp (2019) divide studies of deception into three broad categories. The first group consists of theories based on Grice's (1975) work on cooperation in conversation, such as Information Manipulation Theory which states that deceptive communication is a violation of one or more of Grice's cooperative principles regarding the "quantity, quality, manner or relevance of information" (McCornack, 1992: 2). This also includes Interpersonal Deception Theory, a framework

that expands on Grice's maxims to further consider the interpersonal aspects involved in deceptive communication (Burgoon et al., 1996). The second group focuses on 'speech acts' in terms of intention, for example "lies of commission" when a person is tricked into believing a falsehood, and "lies of omission" when a person is encouraged to continue believing a falsehood (e.g., Bradac (1983) and Vincent and Castelfranchi (1979)). Mett's (1989) assertion that lies consist of three components; "falsification, distortion and omission" also fits into this category. The third group includes work that has sought to develop a "typology of deception strategies" as a means of categorising types of deceptive discourse that go beyond simply stating that something is a 'lie', for instance exaggeration, equivocation, conspiracies, and hoaxes (see Gupta et al. (2012); Hopper and Bell (1984); McGlone and Knapp (2019: 21)). A taxonomy based on examples by Gupta et al. (2012) and McGlone and Knapp (2019) is provided in Table 2-1. Overall, these studies represent interests in defining and mapping the interpersonal aspects of deceptive communication.

Criterion for what constitutes deceptive communication	**Works Referenced**
Violation of Grice's Maxims: Quantity, quality, manner or relevance of information	Barnes (1994) Burgoon et al. (1996) Galasinski (2018) Levine (2014) McCornack (1992)
Speech Acts: lies of commission and lies of omission	Bradac (1983) Chisholm and Feehan (1977) Metts (1989) Turner et al. (1975) Vincent and Castelfranchi (1979)
Typology of Deception Strategies: considering the role of language features such as exaggeration and equivocation	Bavelas et al. (1990) Gupta et al. (2012) Hopper and Bell (1984) McGlone and Knapp (2019)

Table 2-1 - Taxonomy of Models Used to Define Deception Based on Different Criterions

2.3. Fake News Post-2016

This section will discuss the term 'fake news' and how the concept of 'moral panics' can be related to issues regarding fake news and more broadly, deceptive communication. Fake news is treated differently in literature prior to 2016, in comparison to literature published after 2016. Prior to 2016, literature associated fake news as a form of satire (Amarasingam, 2014; Brewer et al., 2013; Reilly, 2012). In comparison, after the events of Brexit and Trump winning the US Presidency, 'fake news' became a term to express "a wider crisis of trust in elites, including political and mainstream media elites, whose members are struggling to maintain their traditional roles in our liberal democracies" (McNair, 2017: xi). Farkas and Schou (2018) refer to fake news as a "floating signifier". By their definition, "fake news" can refer to any of three different types of critiques: a "critique of digital capitalism", "critique of right-wing politics/media" or "critique of liberal/mainstream media" (Farkas and Schou, 2018). Similarly, Wasserman (2017) writes that the definition of fake news differs, depending on how much trust one has in relation to "journalistic and political elites". Therefore, all these definitions of fake news post 2016 share a common theme – a critique of elites and political systems. Thus, 'fake news' can be realised as a politically charged term, in comparison to terms such as 'misinformation' that just reflect a state of information disorder.

Fake news can also be explored through the concept of moral panics. The term 'moral panic' originated as a sociological concept developed by Cohen (1972) to explain the feelings of fear that spread among groups of people when they perceive that something evil is threatening society. Brummette et al. (2018) and Wasserman (2017) have linked fake news to the concept of 'moral panics' in the sense that now anything can be called 'fake news' in order to elicit a strong emotional response. Therefore, the term 'fake news' can be used to discredit an opposing viewpoint, generating a moral panic about specific individuals or organisations that are threatening the integrity of information (Carlson, 2018). In addition, the content of 'fake news' itself plays on the concept of 'moral panics' by triggering fear about "deviant others" that threaten society (Cohen, 1972). Thus, a closer analysis of the linkage between fake news and moral panics in discourse analysis

is needed in order to reveal a more multi-faceted approach to how information disorders are communed around.

The connection between fake news and emotions also relates to several other key terms identified in literature specific to media studies. For example, the notion of "truthiness" created by the talk-show host Stephen Colbert refers to "knowing something by intuition without regard to evidence/facts" and has been discussed by authors such as Boler and Turpin (2008: 386). The "post-truth" era is another related term, defined by d'Ancona as "first and foremost, an emotional phenomenon. It concerns our attitude to truth, rather than truth itself" (d'Ancona, 2017: 126). Terms like 'truthiness' and 'post-truth' also illustrate the impact that emotions have on our interpretation of news and political events. Authors such as Wardle and Derakhshan (2017: 5) have implored people to stop using 'fake news' as term due to its convoluted meanings because it "has also begun to be appropriated by politicians around the world to describe news organizations whose coverage they find disagreeable". The next section will now turn to the information disorder framework as a way of encapsulating the multiple terms that now exist for false information practices.

2.3. Information Disorder

Wardle and Derakhshan (2017) developed a framework for understanding the new complexities of false information in the internet age. The authors specifically use the term "information disorder" as an umbrella term for misinformation, malinformation, and disinformation (Wardle and Derakhshan, 2017), and also as an alternative to simply using the term 'fake news'. The dimensions of information disorder are based on "harm and falseness" (Wardle and Derakhshan, 2017). Looking deeper into information disorder, misinformation is the least severe, referring to misleading information, errors, and satire, all of which inflict the least amount of harm. Disinformation is more severe and refers to intentional deceptive information usually with a political purpose, such as propaganda, information warfare, and astroturfing. 'Malinformation' refers to the sharing of genuine information to cause harm, such as Wikileaks releasing Clinton's emails, or hate speech (Wardle and Derakshan, 2017). The work by Wardle and

Derakhshan (2017) highlights how we should move beyond using the term 'fake news' to define the complexity of deceptive practices online. Figure 2-1 illustrates the distinctions between misinformation, disinformation, and malinformation on a scale of severity. The rest of this section will provide a more detailed explanation of these types of information disorder and how they have been researched as acts of social relations.

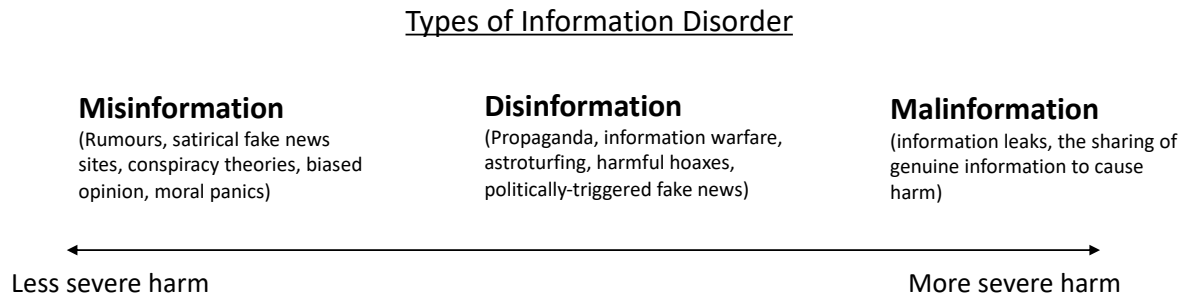


Figure 2-1 - Types of Information Disorder (adapted from Wardle and Derakhshan (2017))

2.3.1. Misinformation

Etymologically, misinformation is the oldest term for false information, originating from the 1580s and meaning the "action of misinforming", with "-mis" referring to "bad/wrong", with a similar lineage to terms such as misappropriate, mistake, misinform, misjudge and misinterpret (Dictionary, nd). In its contemporary meaning, misinformation is defined as "giving false or inaccurate information to" (Dictionaries, 2001). Similarly, "disinformation" is defined in its contemporary sense as "information which is intended to mislead" (Dictionaries, 2001). Thus, from these two similar definitions we can see that in popular discourse people have tended to equate these terms with the same meaning. Nonetheless, these terms actually have quite different histories, as will be further explored.

Studies in misinformation have focused upon issues of rumour (Dare-Edwards, 2014; Zubiaga et al., 2015), gossip (DiFonzo and Bordia, 2007; Foster, 2004, Rosnow and

Foster, 2005), satirical fake news (Boler and Turpin, 2008; Brewer et al., 2013; Dare-Edwards, 2014; Zubiaga et al., 2015), conspiracy theories (Bartlett and Miller, 2010; Mahl et al., 2022), biased opinions (Wallace et al., 2020), and moral panics (Carlson, 2020). As these examples show, misinformation encompasses a broad range of sub-types of communication. Amongst these sub-types, conspiracy theories have been one of the most studied in terms of interpersonal meaning from a range of philosophical, political, and media studies perspectives. The literature on conspiracy theories can be divided into works which focus on conspiracy theories as political propaganda (Cassam, 2019; Yablokov, 2015), versus works that discuss the importance of conspiracies in terms of epistemology (Coady, 2012; Jane and Fleming, 2014; Pigden, 2007). In particular, recent research on conspiracy theories in the COVID-19 era has focused on the power of social relations propagated by social media platforms in persuading people to adopt conspiratorial views (Dow et al., 2021; Restrepo et al., 2021; Van Prooijen et al., 2022).

Rumour and gossip are another misinformation sub-type that have been particularly researched in terms of the enactment of social relations. Even though gossip is an everyday phenomenon, it can be difficult to define since, like casual conversation more generally, it tends to have a prosodic rather than particulate structure, spanning variable stretches of interaction (Eggins and Slade, 2005). It can also be challenging to distinguish gossip from related types of communication such as rumour (Foster, 2004). These problems are exacerbated in online contexts that often involve large datasets spanning multiple types of interactions amongst large volumes of users. Studies of gossip in casual verbal conversations have noted its role in interactively maintaining the values of social groups and have suggested that there is a "dialectical relationship between the linguistic form and the social purpose of gossip" (Eggins and Slade, 2005: 310). In terms of structure, gossip tends to differ from narratives because it does not contain a complication or resolution, instead involving "judgement of an absent other" (Eggins and Slade, 2005: 278). In terms of function, gossiping is commonly seen as an "in-group" activity where users can "air negative opinions without too much sanction" (Robles, 2017: 8). Gossip can also involve the sharing of "extreme opinions" that are used as a strategy for affiliating with others, and as a way of generating "positive politeness" aiding the formation of collective group identity, although sometimes with the risk of

creating discord (Robles, 2017: 7). Accordingly, it is important to analyse gossip from a dialogic perspective, considering the different ways in which users can affiliate or disaffiliate in terms of the social bonds shared in an instance of gossip, as shown by these examples. As this brief overview into the literature on misinformation shows, it is a wide field of research with multiple distinct sub-types of misinformation. Additionally, with the increased focus into researching social media data, the enactment of social relations in the sharing of misinformation have become an increased area of inquiry.

2.3.2. Disinformation

Disinformation originated from the Russian term "dezinformatsiya", (as found in S.I.O.egov's dictionary *Slovar russkogo jazyka* from 1949) (Dictionary, nd). In its contemporary meaning, disinformation has been defined as: "The dissemination of deliberately false information, esp. when supplied by a government or its agent to a foreign power or to the media, with the intention of influencing the policies or opinions of those who receive it" (Dictionaries, 2001). In addition, in both journalistic and academic discourse, a strong connection has been made with disinformation and Russian interference, as even the term disinformation originated from the Cold War era (Haiden and Althuis, 2018). As Karlova and Fisher (2013: 4) write: "the strong association between disinformation and negative, malicious intent probably developed as a result of Stalinist information control policies". Presently, organizations such as the *European External Action Service East Stratcom Task Form*, *Eu vs. Disinfo,* and *Stop Fake*, have all used the term "disinformation" in order to address false information particularly spread by Russian organizations. Many academic articles have also been written relating Russia to disinformation (Bennett and Livingston, 2018; Mejias and Vokuev, 2017; Yablokov, 2015). Research has been conducted into how Russian propaganda promotes conspiratorial thinking (Starbird et al., 2019) and how Russian narratives have been supported by organisations and individuals belonging to the far-right, who promote anti-Semitism and Islamophobia (Culloty et al., 2020; Starbird et al., 2019). Thus, the word 'disinformation' continues to hold strong connections to Russia, even if the definition of disinformation does not explicitly state this.

There is also a strong connection between disinformation and propaganda which from 1929 onwards has referred to "material or information propagated to advance a cause, etc." (Etymology dictionary, nd). For example, Søe (2018) defines propaganda as "information of a biased/misleading nature to promote a political cause". Biddle in a *Psychological Definition of Propaganda* writes how propaganda has relied on methods including "persuasion", "direct emotional appeal", "direct suggestion" and "indirect suggestion" (Biddle, 1931: 283). In addition, he writes that "the emotional pattern of "we" versus an "enemy" runs through all propaganda, of war or of peace" (Biddle, 1931: 285). This is similar to how disinformation campaigns rely on an enemy such as Russia vs. the West (Lucas, 2014; Mejias and Vokuev, 2017; Polyakova and Boyer, 2018). Thus, the solution to combating 'disinformation' becomes getting rid of a single enemy. Therefore, there are many similarities with propaganda and disinformation. In this case, information manipulation and ideologies can also be considered. van Dijk (2006: 359) writes that manipulation consists of "social power abuse, cognitive mind control and discursive interaction". These qualities can be linked to deceptive communication and the intentional harm caused by disinformation. As we can see from this discussion of the definitions of disinformation and propaganda, disinformation is the broader umbrella term referring to any false information that has been deliberately created, whilst propaganda refers to deliberate false information shared to promote a political cause. According to the information disorder framework, disinformation can also incorporate astroturfing, in which people pretend to be ordinary citizens in order to influence political behaviour (Keller et al., 2020; Zerback et al., 2021), as well as, harmful hoaxes (Finneman and Thomas, 2018; Kumar et al., 2016; Sellnow et al., 2019), and politically-triggered fake news (Bakir and McStay, 2018). All these examples highlight the complex sub-types within disinformation.

2.3.3. Malinformation

Malinformation causes the most harm according to the information disorder framework. It refers to the sharing of genuine information, or information based on reality, to cause harm such as the leaking of information that may place people in danger, or racist and white supremacist discourses that can place individual's lives in

danger. It should be noted that hate speech can be placed in all the sub-types of misinformation, disinformation, and malinformation in terms of the information disorder spectrum, depending on its severity, and whether it represents deliberately created material to cause harm, or manipulates real information in order to place someone's life in danger.

For example, issues of hate speech and white supremacy have manifested in different ways online, merging with conspiracy theories and deceptive disinformation practices. White supremacy online is connected to platformed racism, as people with white supremacist values mobilise online, centring their whiteness (Western values) and targeting hatred at a racialized 'other' (Daniels, 2009; Gillborn, 2006). This white supremacy is often manifested in the form of conspiratorial discourse. Conspiracy theories have been linked to racist and xenophobic discourses, for example, in the form of islamophobia (Farkas et al., 2018), or antisemitism (Allington et al., 2021). Research has also been conducted at the intersection of mediatization and hateful language. In particular, conspiracy theories have been analysed as sociocultural changes developed in relation to mediatization. For example, conspiracy theories have been theorised as no longer belonging to the "fringes of media" but rather as part of the established news narrative (Konkes and Lester, 2017) or theorised as contributing to the mediatization of conflict by government agencies (Culloty et al., 2020).

Mediatization and hate speech have also been explored in relation to the "politicization" of the refugee crisis in Europe (Krzyżanowski et al., 2018; Krzyżanowski, 2018) particularly focusing on the discursive shifts in the mediatization of right-wing populist parties, who write strongly ideological messages on social media to create "the image of dialogue with citizens and other strands of the public sphere" (Krzyżanowski, 2018: 79). As these examples from the literature show, mediatization and hateful language have been the focus of detailed analysis. However, as Lim (2020: 606) identifies, media scholars need to connect with other disciplines in order to address the "challenges of digitalization and mediatization" in relation to the rise of hate speech via social media networks. By engaging with the methods of multiple disciplines regarding these

overarching issues, "academically informed, evidence-based, and finely balanced" results can be achieved (Lim, 2020: 606).

2.3.4. Issues in Defining Information Disorders

Presently, academic research has been divided when considering the linkage between deception and lying to literature on misinformation and disinformation. As previously noted, there is a lot of literature that has considered the definitions of deception and lying, in relation to verbal and non-verbal communication. In media and communication studies, misinformation, disinformation, and 'fake news' have been extensively written about, even if the definitions of these terms remain contentious. However, the exact connection between deception and lying, in relation to misinformation and disinformation has not been extensively explored. Perhaps, this is due to the relative contemporaneity of the terms misinformation and disinformation, and that research into the deception and lying is outdated in the sense that it has not been fully explained for the internet age. Therefore, this section of the literature review hopes to bridge studies in deception and lying, to studies in misinformation and disinformation, and outline some potential issues in defining and discussing information disorders.

Søe (2018) is one of the few authors that discuss misinformation and disinformation in relation to literature on lying and deceiving. As Søe (2018: 14) writes:

> "The treatment of misinformation and disinformation as two distinct notions is in line with the general conception within the literature on – lying, misleading, and deceiving that it is necessary to distinguish between lies – (believed-false statements), misleadingness (based on inaccuracies or implicatures both verbal and gestural), and deception (successful and intentional misleading and lying) (cf. Mahon, 2008)."

There is a need to link studies of deception to studies of misinformation and disinformation. However, as previously noted, minimal connections are made in linking

deceptive communication literature to the issues reflected in media studies surrounding fake news and information disorder. Søe (2018: 311) also highlights this when they pose the question: "what characterizes deception as opposed to non-deception?", arguing that we must work out if misinformation and disinformation both contain deceptive clues, or if only one dimension contains deceptive cues. Although Søe (2018) does link deception to mis/disinformation, in this case they are not treating deception as the overlying superordinate:

> "If, for instance, both misinformation and disinformation have deception as one of their features, then deception is not a sufficient feature for the algorithms to detect –the algorithms should also look for something else. If, however, deception is only a feature of disinformation then the algorithms have to detect for deception." (Søe, 2018: 311).

Instead, deception seems to be attached to the question of intention, which is difficult to determine through textual analysis, as intention is inherently psychological. Thus, in order to expand Soe's (2018) notion of deception, considering mis/disinformation under the umbrella term of information disorder seems more plausible. Although misinformation may be an unintentional mistake, in the social media sphere, incorrect information can still have disastrous effects. Thus, one is still being deceived when presented with cases of misinformation. As we do not know one's intentions on social media (in the sense that we cannot read people's minds) this means that mis/disinformation should both be considered as information disorder.

The distinction between the terms 'misinformation' and 'disinformation' has been a point of contention in academic literature and popular culture. Some authors use these terms interchangeably, not paying much attention to the distinctions, others use one term only, believing that it also encompasses the other, whilst another group of authors have made attempts to develop a typology of definitions clearly distinguishing misinformation from disinformation. In addition, disinformation has been focused on much more than misinformation, despite both these forms of information disorder being prevalent.

Not all works of academic literature have drawn a distinction between "misinformation" and "disinformation". For example, consider the terms used by McNair (2017: 13) in the following:

> "In the UK, too, the Leave referendum generated many examples of deliberate misinformation and disinformation".

In this example, misinformation and disinformation are both treated as "deliberate" and yet according to most common definitions of misinformation, the term refers to an "honest, unintentional mistake". Zhou and Zhang (2007: 806) also do not distinguish between misinformation and disinformation and therefore describe misinformation as "concealment, ambivalence, distortion and falsification", and ignore the more accepted definition of misinformation as an unintentional error. Fallis writes that "inaccurate information (or misinformation) can mislead people whether it results from an honest mistake, negligence, unconscious bias, or (as in the case of disinformation) intentional deception" (Fallis, 2014: 136). Thus, Fallis defines information, misinformation, and disinformation as representational content, where it is further specified that misinformation is inaccurate and misleading, and disinformation is intentionally or non-accidentally misleading. This misleadingness is often generated through Gricean implicatures, in terms of quality violations – especially in instances of disinformation.

Similar to the way in which authors such as McCornack (1992) have defined deceptive discourse in relation to Grice's maxims, Søe (2018) provides a Gricean perspective on misinformation and disinformation. Søe (2018) defines misinformation as "unintended misleading" whilst disinformation refers to "intentional deception". Thus, her criterion for defining these terms is situated around intentionality and misleadingness – two concepts that also shape how mis-information/dis-information are distinct from 'information' that is supposed to be intentional and non-misleading. Søe (2018) bases these definitions on a review of authors who have written about misinformation (Fox, 1983), disinformation (Fallis, 2015) and semantic information (Floridi, 1996). Søe (2018) also provides some linkage between defining lies, deception, and

misinformation/disinformation. Lies are defined as "believed false statements" and deception as "successful and intentional misleading and lying" (Søe, 2018). A third term "misleadingness" is also added to refer to "inaccuracies and implicatures" (Søe, 2018). As this section has shown, there is a lack of consensus in academic research regarding how information disorders are named and categorised, due to this confluence of terms, this thesis adopts the umbrella term of information disorder in order to reflect this complex array of research.

2.4. Algorithmic Approaches to Understanding Information Disorders

Previous approaches to detecting information disorders have suggested methods such as fact-checking websites (e.g. *Politifact, Buzzfeed News* and *Snopes*), computational processes (*PHEME Project*), increasing information literacy, and applying psychological studies. These approaches can be divided into those aiming to deliver real-time solutions for detecting deceptive communication so that websites can flag it as it appears, and approaches with longer-term goals such as educating the general populace about information literacy and funding fact-checking organisations. There are also solutions that attempt to merge human fact-checking and algorithmic detection methods together, some of which will be further explained in this section.

Rumour detection has been a key focus of the development of algorithms seeking to differentiate real from false information on social media. Rumours can be considered as a form of deceptive communication. Additionally, rumours fit the category of misinformation (unintentionally spreading false information) or the category of disinformation (deliberately spreading rumours in order to incite fear and confusion). Different approaches have been applied to this dilemma, either broadly attempting to detect rumours (intentional or unintentional) or delving deeper into these definitional areas of concern. Table 2-2 outlines some of these key projects, considering how these projects have defined mis/disinformation in similarity to work by Søe (2018).

Name of Detection Project	Methodology	How Misinformation and Disinformation is Defined
PHEME Project (Zubiaga et al., 2017)	A combination of manual annotation and computational processes	No distinction between mis/disinformation, the focus is on rumour detection.
Hoaxy (Shao et al., 2016: 745)	All algorithmic detection. Only tweets containing hyperlinks are analysed and cross-checked with fact-checking websites.	The Hoaxy platform defines misinformation as the "unintentional spread of false or inaccurate information" (Shao et al., 2016: 745). There is no mention of disinformation.
Diffusion Model (Karlova and Fisher, 2013)	Theoretical approach, providing a model for future detection projects. No empirical data collected currently.	The Diffusion model distinguishes misinformation as "inaccurate information" and disinformation as "deceptive information" (Karlova and Fisher, 2013).
Cognitive Psychology (Kumar and Geethakumari, 2014)	Model is based on a cognitive psychology approach and algorithmic process. Only retweets are collected.	Misinformation refers to "false or inaccurate information", especially that which is "deliberately intended to deceive" (Kumar and Geethakumari, 2014: 3). Disinformation is "false information that is intended to mislead" (Kumar and Geethakumari, 2014: 3) and refers to a "sophisticated deceit process".

Table 2-2 -Mis/Disinformation Detection Projects

Therefore, based on these projects, we can see that a lot of work has already been completed in terms of detecting true versus false statements. Although some of these projects attempt to distinguish misinformation versus disinformation, they have not achieved this by using linguistic criteria or by a consideration of evaluative meanings in a text. For example, sentiment analysis has automated emotive features of language (in

terms of affect) but has not considered the linkage to mis/disinformation and other evaluative meanings such as the judgement of someone or appreciation of an object.

Previously, rumours have been studied from fields such as psychology (DiFonzo and Bordia, 2007; Rosnow and Foster, 2005), celebrity studies (Dare-Edwards, 2014) and journalism studies (Hermida, 2010; Mintz, 2012). These studies have all been active in defining rumours through case studies and commentary on current events. Nonetheless, these studies have not attempted to apply a methodology to distinguishing rumours from non-rumours online.

Currently, the study of distinguishing rumours from non-rumours on Twitter, has been an area of focus in the fields of computational linguistics and social informatics. Giasemidis et al. (2016) have analysed the veracity of twitter rumours by the application of "autonomous message-classifiers" that can filter relevant and trustworthy information from twitter. Similarly, machine-learning techniques have also been applied by Gupta et al. (2014) with the development of the machine learning algorithm 'TweetCred' to understand the credibility of a tweet.

Understanding the nature of rumours on twitter through computational linguistic methodologies, has been applied to work in journalism and disaster mitigation. In particular, Lukasik et al. (2015), Procter et al. (2013) and Vis (2013) have focused on understanding tweets from the 2011 England riots. Jin et al. (2014) have applied a quantitative analysis of tweets during the Ebola Crisis and Mendoza et al. (2010) have applied an aggregated data analysis towards understanding patterns of rumour propagation based on tweets from the 2010 earthquake in Chile. All of these studies have focused on using multi-task learning and computational linguistics to identify keywords or patterns associated with each rumour but have not developed methodologies for distinguishing rumours from non-rumours. Other approaches to detecting fake news and satire have similarly applied computational processes. Rubin et al. (2016) have created machine-learning based algorithms to develop 'satirical cues' to detect fake news. Shao et al. (2016) and Shao et al. (2017) have developed "Hoaxy", a machine learning algorithm, that can track online misinformation.

Approaches that combine manual annotation and computational linguistics have also been applied to the study of rumours and misinformation. Volkova et al. (2017) have incorporated neural network models and linguistic annotations to detect "satire, hoaxes, clickbait and propaganda". Qazvinian et al. (2011) applied a mixed method approach – using manual sentiment analysis and computational processes to understand misinformation in relation to microblogs and detect "disinformers". The manual sentiment analysis conducted, consisted of annotating posts as "rumour, non-rumour, believe or deny" (Qazvinian et al., 2011). This provides an overview of algorithmic approaches to understanding information disorders. As these examples have shown, the focus of these studies is on detecting and automating isolated linguistic cues in specific instances of communication rather than considering the interpersonal dimension of social bonds that underline these communications.

2.5. Discourse-Analytic Approaches to Understanding Information Disorders

A range of discourse-analytic approaches to understanding information disorders have been explored, particularly in the sub-fields of forensic linguistics and corpus linguistics. Firstly, it is important to clarify the meaning of a discourse-analytic approach as the term discourse itself can be "fuzzy" (Van Dijk, 1997: 1) and depending on one's academic discipline, the meaning of discourse analysis can vary. In common sense terms, discourse refers to a form of language use or ideas around a concept, however, in academic terms, discourse is more theoretical and embodies the functional aspect of the communicative event (Van Dijk, 1997). In other words, the idea comprising that "people use language in order to communicate ideas or beliefs….and do so as part of more complex social events" (Van Dijk, 1997: 1). Discourse analysis approaches vary, encompassing at least four main types of work: speech act theory and systematic accounts of conversational exchanges, psychologically orientated discourse analysis, sociology of scientific knowledge, and approaches drawing on continental social philosophy and cultural analysis (Potter and Wetherell, 2002: 47).

The majority of linguistic research into information disorders focuses on a forensic linguistics perspective. Forensic linguistics has been concerned with issues such as how

forensic authorship analysis and evidencing, can be used in cases of fraud, plagiarism, and criminal courts (Gibbons and Turell, 2008). The analysis of scam emails has also been another area of focus in linguistics by authors such as Ott et al. (2011) and the Scamseek project by Patrick (2006). Corpus linguistic approaches have been a particular focus of research into information disorders. This includes research into trolling (Hardaker, 2010), abusive language (Clarke and Grieve, 2017), and conflict about vaccination (Coltman-Patel et al., 2022). These approaches, whilst successful in determining patterns of deceptive cues in language, do not examine the key social bonds particular to users, a task that is well suited to a social semiotic approach grounded in affiliation.

The discourse analyst Galasinski provides a comprehensive analysis of issues regarding how one can analyse deception linguistically. Galasinski (2018: 522) introduces the term "all-linguistic" research into deception, which means "that both the deceptive communication (the 'lie') and the reality it misrepresents are linguistic". In addition, they discuss the issue of intention regarding how one can research deception linguistically:

> Yet, despite lack of insight into the deceptive intention, one has to solve two problems. First, access both to misrepresentation or deception and to the reality it misrepresents. Second, such data allow the discourse analyst to focus on what s/he can analyse-real life discursive data that are amenable to analysis with discourse-analytic tools. (Galasinski, 2018: 522).

Galasinksi's solution for the discourse analyst to not forget the importance of social context: "The ideal data for the discourse analyst consist in two representations: First, the initial description of reality and then, second, the misrepresentation of that description" Galasinski, 2018: 522). Rather than looking at just one instance of deceptive communication, the analyst needs a large collection of data from which they can derive the actual context of the situation and find evidence for why something is a misrepresentation. Thus, Galasinksi effectively addresses the question of how one can understand how intention unfolds in deceptive discourse textually. This thesis will

apply a similar philosophy to understanding information disorders, relying on a large collection of data to make interferences about what is real or false.

Discourse analyses of white supremacy and conspiracy theories in digital communication have also been an area of examination, for example, in relation to the psychological constructs of tweets by political extremists and conspiracists (Fong et al., 2021), the discourse of tweets by alt-right supporters (Panizo-LLedot et al., 2019), and the narrative frameworks of posts by conspiracists on social media forums (Shahsavari et al., 2020). The majority of these linguistic analyses are focused upon automated analyses of keywords due to the large datasets these researchers are working with. Whilst this research is valuable, a focus purely on identifying keywords neglects the broader discourse semantic functions of conspiratorial language. The approach in this thesis goes beyond identifying specific keywords associated with these creators and communities, and instead identifies broader discourse semantic functions to illustrate how these communities are addressed and promoted.

Understanding the importance of evaluative language and stance in white supremacist and conspiratorial discourse is also an increased area of focus by researchers working across a range of discourse analysis perspectives (Demata et al., 2022; Szenes, 2021). The aims of this thesis relate to previous studies on the mediatization of societal threats (Krzyżanowski et al., 2018) by highlighting the power of rhetoric in 'othering' certain groups in order to gain popularity and emphasising how 'the politics of fear' (Wodak, 2015) works at a discourse semantic level. Therefore, this thesis engages with previous valuable research into information disorders and societal threats but includes the added dimension of affiliation in order to emphasize how social values are bonded around. This increased interest in discourse-analytic approaches to understanding information disorders demonstrates a need to further delineate hateful communities according to these linguistic strategies and collaboratively design materials that educate a diverse range of populations.

2.6. Multimodal Approaches to Understanding Information Disorders

Current work has begun to consider how information disorders are related to visual images and videos. A multimodal approach, that is, an approach that incorporates the analysis of modes of communication beyond language, such as image and gesture, has accompanied textual analysis. Although images and videos are such a significant part of information disorders on social media, not many projects have taken on the task of understanding deceptive images and videos. The few studies that exist include Gupta et al. (2013); Glenski et al. (2019) and Lu et al. (2008). However, these studies come from the field of computer science studies, and do not take a multimodal social semiotic approach to analysing images and videos. Thus, more work that analyses the social semiotic aspects of multimodal texts engaging with a wide range of information disorders is needed. Work that has taken a multimodal approach to studying certain aspects of information disorders includes research into the relationships between memes and fake news (Smith, 2019), detecting fake news stories using multimodal methods (Singh et al., 2021), and combining multimodal discourse analysis methods with natural language processing to analyse tactics in violent extremism (O'Halloran et al., 2019; Wignell et al., 2021).

In terms of research into the role of visual images such as screenshots in the propagation of manipulative information, 'evidence collages' in media manipulation campaigns have been examined from an ethnographic perspective (Krafft and Donovan, 2020). These collages incorporate screenshots as a form of visual evidence that function as "a key strategic element in the formation and spread of disinformation" (Krafft and Donovan, 2020: 205). Screenshots have also been studied as technologies for public shaming. For example, a qualitative thematic analysis of news media articles about the cases of Amanda Todd, a Canadian teenager who was cyberbullied and blackmailed with screenshots taken of her without her consent leading to her taking her own life, and Anthony Weiner, a US Congressman whose political career ended when screenshots were revealed of his extramarital flirtations, highlighted how screenshots are entangled between the notions of permanence and ephemerality, and the boundary violations that occur when the private is made public (Corry, 2021). Some studies have noted the role

of screenshots in the propagation and visibility of racist discourses where "platform collapse" and "mediated spillover" occurs, as screenshots are shared across various social media platforms without users understanding their proper context (Bigman et al., 2022: 4). Whilst there is some positive potential for screenshots to call out social media users who engage in a "tweet and delete" culture of harassment, this needs to be weighed against their potential negative impact in proliferating hateful content, for instance, by amplifying racist content (Bigman et al., 2022: 11). On the other hand, screenshots may act as a tool for sousveillance, an effective form of visual persuasion, whereby organizations such as Racism Watchdog draw attention to online injustices (Jenkins and Cramer, 2022). These examples show how multimodal approaches to information disorders have become an increasing area of interest in academic research.

2.7. YouTube as a Social Network Propagating Information Disorders

The sharing and proliferation of misinformation and disinformation have become a growing concern in an age of social media platforms which are increasingly developing into incubators for hateful discourse. Platforms such as YouTube are amplifiers and manufacturers of hateful discourse due to the affordances, business models and cultures that they maintain through forms of 'platformed racism' (Matamoros-Fernández, 2017). YouTube has become a social media platform displaying acts of conflict and antagonism (Burgess and Green, 2018; Pihlaja, 2014). It has also been used for spreading disinformation and extremist views, ranging from terrorist recruitment (Andre, 2012; Klausen et al., 2012), state propaganda (Golovchenko et al., 2020), and right-wing extremism (Ekman, 2014; Levy, 2020; Lewis, 2018), concepts that will be further explored in this section.

Previous qualitative research regarding deceptive content on YouTube, has explored right-wing extremist communities (Ekman, 2014; Levy, 2020 Lewis, 2018), racist influencers (Johns, 2017; Murthy and Sharma, 2019), populist YouTubers (Finlayson, 2020; Zuk and Zuk, 2020) and conspiracy videos (Allington and Joshi, 2020; Mohammed, 2019; Paolillo, 2018). Critical discourse analysis of YouTube videos has been undertaken in order to understand discourses of xenophobia (Asakitikpi and Gadzikwa, 2020),

xenophobia and misogyny (Kopytowska, 2021) and racism (Hokka, 2021). Thus, a wide array of different communities and methodologies have already been researched in relation to deceptive content on YouTube. YouTube comments in particular are important as an object of study in studies on information disorder because they provide a wealth of freely available public data, that stretches across "international and inter-generational audiences" (Thelwall, 2018: 304). In addition, studying YouTube comments allows us to understand the participatory framework of YouTube as a media platform that enables interactions in multiple ways (Dynel, 2014). For example, YouTube comments can serve as a space to pseudonymously provide commentary on an issue addressed in a video but avoid directly engaging with other user comments. Or conversely, YouTube comments are a space to directly reply to and name users in interactions that extend beyond the content of the video. Studying YouTube comments gives us the ability to conduct a close linguistic analysis on clearly defined historical moments, an important criterion if wanting to understand how information disorders are mobilised in society.

Whilst issues of hate speech and disinformation on social media have been studied widely, there is a need to move beyond solely researching platforms such as Twitter that have currently dominated disinformation and racism research (Matamoros-Fernández and Farkas, 2021). In comparison to Twitter, YouTube is a less studied platform but one that needs further exploration, particularly in regards to ensuring platform diversity and cross-platform analyses (Matamoros-Fernández and Farkas, 2021). The unequal amount of research into YouTube as a platform could be attributed to the methodological challenges of analysing YouTube as a multimodal platform incorporating many different forms of data such as audio-visual content, transcripts, and comments. YouTube also does not lend itself well to a quantitative analysis due to these multiple features, meaning that it is difficult to understand its impact and user engagements via numerical data alone. The quantitative studies that do exist on YouTube and are mainly focused on analysing the metadata of YouTube videos or sentiment analysis of content as positive or negative (Donzelli et al., 2018), but have not applied a closer inspection of comments or audio-visual data.

Another issue problematising the use of YouTube to study information disorders is the difficulty of analysing YouTube's algorithms. YouTube has never shared details of its specific algorithms. Some research such as work by Rieder et al. (2018) and Bishop (2018) has sought to uncover YouTube's algorithms through methods such as rank visualisations (the visualisation of YouTube search results) and reverse engineering. Nevertheless, algorithms should not be seen as the sole determinant of how deceptive or problematic communication spreads. Deceptive communication is not a purely technical problem but "also a fundamentally social problem" (Lewis, 2018). The qualitative approach of this thesis addresses deception and information disorders as key social problems evident on YouTube and develops tools for understanding information disorders on YouTube in terms of social bonding.

Overall, information disorders on YouTube, encompassing issues like internet hoaxes, white supremacy and conspiracy theories are pertinent issues to be studied, requiring both qualitative and quantitative approaches to identify and categorise false, racist or conspiratorial discourse, in order to understand and mitigate these issues. Despite the important work already carried out in this space, there has yet to be a concentration on identifying via linguistic and multimodal methods the strategies YouTubers use to form social bonds in order to affiliate with others, nor studies of how these social bonds are legitimated in language and through visual resources.

2.8. Conclusion

This chapter has explored the literature regarding deceiving and lying, and deceptive communication, linking these to current approaches to analysing information disorders on social media. As has been highlighted, many of these studies have focused on distinguishing viral rumours from non-rumours. They have generally not considered how analysis of evaluative language can contribute to understanding the social bonds at stake in the construction of information disorders in their various forms. This thesis aims to fill this gap by applying a social semiotic approach to online information disorders by exploring how interpersonal meaning and social relations are linguistically enacted in these contexts. The SFL method of this thesis provides a way of delving

deeper into the social values of conspiratorial communities identified by computational social science studies (Shahsavari et al., 2020) by highlighting the affiliation strategies used to make these social values prominent. Methods in SFL also intersect with the key issues raised in communication and media literacy studies as explored in this chapter, that highlight the complexity in effectively communicating the multi-layered issues at hand and translating academic results to varied audiences.

3. CHAPTER THREE THEORETICAL FRAMEWORK: A DISCOURSE-ANALYTIC APPROACH GROUNDED IN SOCIAL SEMIOTICS

3.1. Introduction

This chapter details the main theoretical framework which underlies this thesis – a social semiotic approach informed by Systemic Functional Linguistics (SFL) and Multimodal Discourse Analysis (MDA). It begins with an introduction to SFL as a linguistic model and then explains the main analytical methods that this thesis draws from SFL. The second part of this chapter provide an overview of MDA. It focuses on multimodality in relation to legitimation, identity, computational studies, and the analysis of social values. By outlining the significance of previous academic work on these issues within SFL and MDA, and how these theories can be practically applied to issues concerning the study of information disorders, this chapter will illuminate the theoretical foundations of this thesis.

3.2. Systemic Functional Linguistics

Systemic Functional Linguistics (SFL) studies meaning-making resources in language in terms of their social context (Halliday, 1978: 122). SFL considers language as both a functional process, as its function is to make meanings, and a semiotic process, that is, a process of making meanings via a selection of choices. SFL is systemic in the sense that it applies an "analytical methodology which permits the detailed and systematic description of language patterns" (Eggins, 2004: 23). Additionally, SFL uses 'system networks' as a graphical formalism to describe the sets of options available for a chosen entry condition. System networks represent the paradigmatic organisation of language, in other words, meaning arising from choice (Matthiessen et al., 2010). An example of a system network is illustrated in *Figure 3-1*. The curly bracket represents simultaneous systems of 'and', whilst a square bracket represents parallel systems of 'and' where only

one feature can be selected. It is this systematic approach to language that is characteristic of SFL.

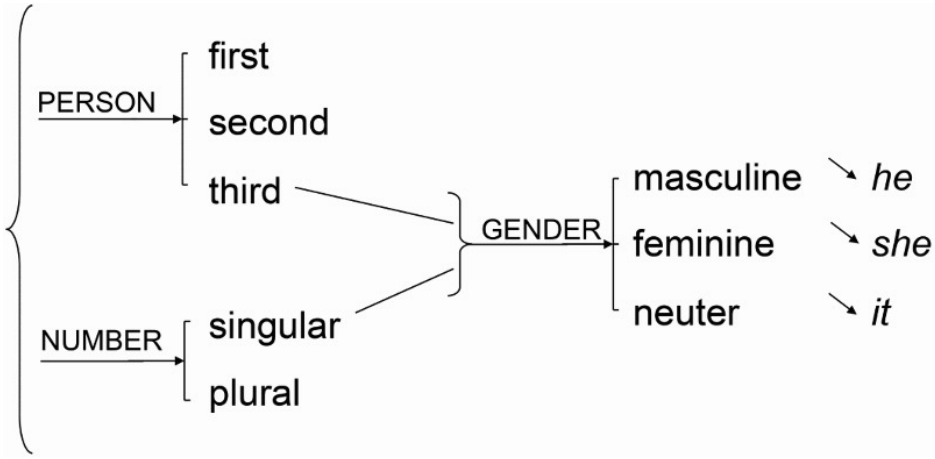


Figure 3-1 - Example of a System Network (from Martin, 2016)

The distinct linguistic approach of SFL was developed by the British linguist Michael Halliday from the 1960s onwards (Halliday, 1961). Halliday's approach was inspired by a range of theories that foregrounded the social significance of language. Examples of these theories include, Saussure's notion of language as a system of signs (De Saussure, 1916; Holdcroft, 1991), Hjelmslev's theory of language as a stratified systems of signs (Hjelmslev, 2019; Taverniers, 2011), Firth's notion of systems and structures in language (Firth, 1957; Firth, 1950), Bernstein's model of language as codes (Bernstein, 2003), and Malinowski's work on the 'context of situation' (Malinowski, 2001; Robins, 2004). Halliday also took inspiration from the Prague School of Linguistics (Vachek, 1966), from authors such as Danes (1964), Firbas (1964) and Vachek (1964) that focused on developing a structural literary analysis (Martin, 2016). SFL is an 'appliable linguistics' which Halliday described as:

> "…capable of being applied to the problems, that are being faced all the time by the many groups of people in our modern society who are in some way or other having to engage with language" (Halliday, 2008: 7).

A core concern of SFL is developing an approach to the study of language that can be applied to many different practical situations and areas of research in order to address particular social issues and problems adhering to Halliday's notion of an appliable linguistics. For example, SFL has been applied to research in a diverse range of fields such as language education (Christie, 1999; Christie and Martin, 2009; Unsworth, 2005), academic writing (Hood, 2010; Stenglin and Cléirigh, 2020), health (Karimi et al., 2018; Matthiessen, 2013; McDonald and Woodward-Kron, 2016), social media (Zappavigna, 2012; Zappavigna, 2018), casual conversation (Eggins and Slade, 2005), computational contexts (Bateman et al., 2019; Teich, 1999), media discourse (Feez et al., 2008; White, 1998), history (Martin and Wodak, 2003), administrative language (Iedema, 2003), business (Szenes, 2017), and physics (Doran, 2017). Additionally, SFL has expanded to semiotic modes other than language, which will be explored in more detail in Section 3.3 on Multimodal Discourse Analysis (MDA).

A key insight of SFL is that language simultaneously enacts three main functions, referred to as metafunctions (Halliday and Matthiessen, 2013; Martin, 1992). These metafunctions are the *ideational* that construes experience, the *interpersonal* focusing on relationships, and the *textual* that is based around organisation (Halliday and Matthiessen, 2013). Research into information disorders has particularly neglected the interpersonal metafunction (in terms of how people relate to or bond around deceptive or problematic ideas and values). Instead, research has tended to focus on the ideational dimension of information disorders (what is experienced as true or false). Therefore, an SFL approach can provide a foundation to understanding information disorders, and its interpersonal significance.

SFL maps out context and language according to the hierarchy of stratification. This means context and language are organised in patterns of patterns (Martin, 2016) in terms of abstraction with higher order strata realised by the stratum below (Figure 3-2) As shown in Figure 3-2 the strata span the following:

- **Genre**: In Halliday's and Hasan's work genre is defined as the register of language or the mode variable of context (Halliday and Hasan, 1989; Matthiessen

et al., 2010) and was not officially part of the hierarchy of stratification. However, overtime genre has evolved in SFL to consider genre as part of the hierarchy of stratification, or in other words, the model of context. Martin and Rose define genre as "staged, goal oriented social processes" and as a "recurrent configuration of meanings that enact the social practices of a given culture" (Martin and Rose, 2008: 5).

- **Register**: Register theory describes "the impact of dimensions of the immediate context of situation of a language event on the way language is used" (Eggins, 2004: 9). These key dimensions are field (what is happening with the social action taking place i.e., the topic or focus of the activity), tenor (who is taking part in language i.e., the relations of power and solidarity between people) and mode (what role language is playing i.e., the organisation of the text) (Eggins, 2004; Martin and Rose, 2008).

- **Discourse semantics**: In SFL, discourse semantics is concerned with what is between the analysis of grammar and the analysis of social activity, in other words, texts that are "bigger than a clause and smaller than a culture" (Martin and Rose, 2003: 3). The discourse analyst "employs the tools of grammarians to identify the words in passages of text, and employs the tools of social theorists to explain why they make the meanings they do" (Martin and Rose, 2003: 4). It is at the level of discourse semantics that this thesis is most concerned with.

- **Lexicogrammar**: This refers to the resources for "construing meanings as wordings – the combination of grammar and lexis (vocabulary)" (Matthiessen et al., 2010: 131).

- **Phonology**: This is the last stratum within the expression plane of language. Phonology encompasses phonetics (the human articulatory and auditory systems) and graphology (the study of written language) (Cléirigh, 1998; Martin, 2016; Zappavigna et al., 2010).

All these different strata are related to each other through realisation and metaredundancy i.e. "patterns of patterns" (Martin and Rose, 2008: 27) or the "redundancy of redundancy" (Matthiessen et al., 2010: 21). For example, genre is realised by register, and register is realised by discourse semantics (Martin and Rose, 2008). Thus, realisation refers to the relationship between different strata. The relationship between these forms of strata is shown in *Figure 3-2*. In understanding these different strata, we must also consider the concept of 'rank' – "the hierarchy of units based on composition" (Matthiessen et al., 2010: 21). For example, there are two key rank scales, the clause-group/phrase-word-morpheme scale that links to the cline from genre, register, discourse semantics and lexicogrammar, and the tone group-foot-syllable-phoneme scale that is linked to phonology (Matthiessen et al., 2010), illustrated in *Figure 3-3*. Thus, within these different strata there are different features of language that are explored.

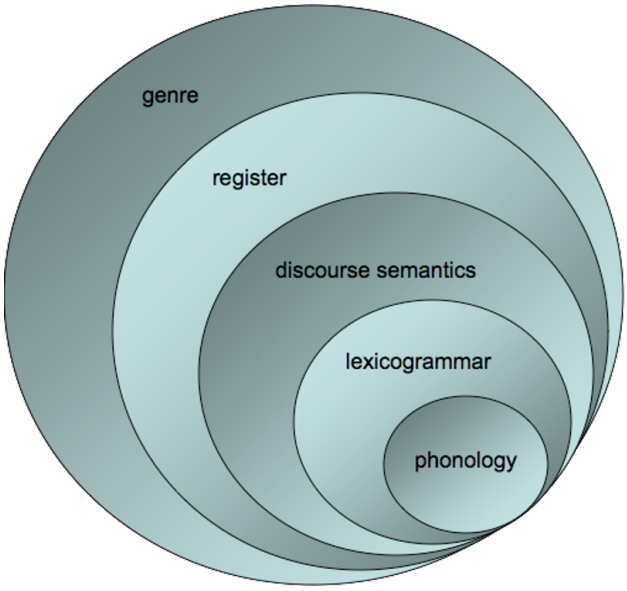


Figure 3-2 - SFL Stratum of Meaning (from Martin 1992)

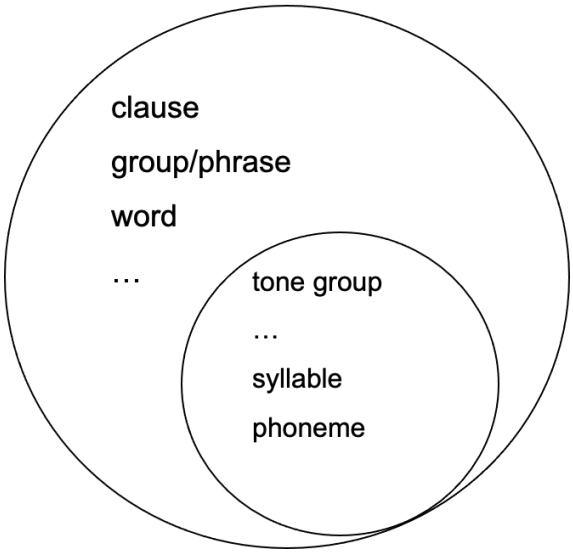


Figure 3-3 - Ranks of Meaning in SFL (from Martin, 1992)

There are two other key concepts essential to providing an overview of SFL: individuation and instantiation. Individuation is organised according to the cline from the collective to the individual and will be further explored in section 3.2.5 (Knight, 2010a; Matthiessen et al., 2010). To summarise very briefly, individuation refers to meaning potential according to people (the users of language) rather than a focus on the systems or uses of language itself. It shows the relations between reservoir (the system of meanings available in a culture) and repertoire (the meanings available to an individual). Instantiation is related to the cline of individuation and refers to the "overall meaning potential of a language and the cultural meaning potential that it operates in" (Matthiessen et al., 2010: 117). In other words, it is focused on the relationship between an instance of language and the system that lies behind it (Matthiessen et al., 2010) in contrast to being concerned with users of language as with individuation. It shows the relations between a language system and an instance of a text. A diagram of how realisation, instantiation and individuation work together in the theory of SFL is shown in *Figure 3-4*. We will now discuss different theories within SFL that will be the main focus of this thesis.

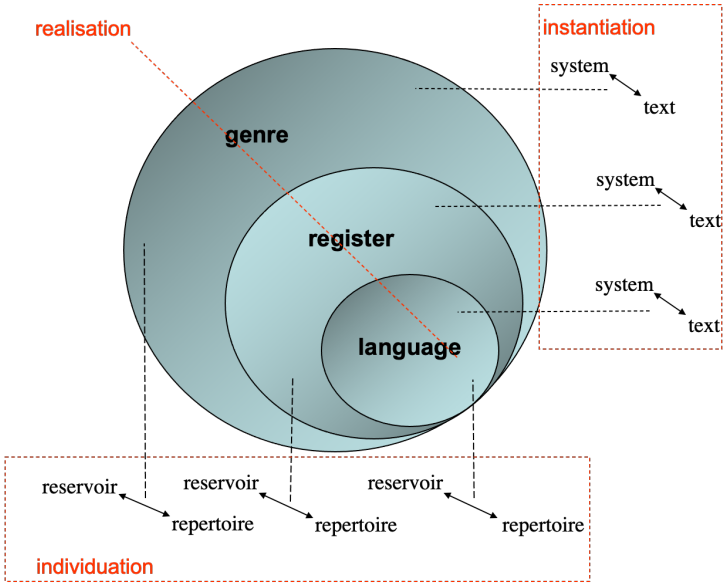


Figure 3-4 – Relationship between Realisation, Instantiation, and Individuation Hierarchies (from Martin 1992)

3.2.1. Appraisal

The appraisal framework systematises evaluative meaning at the level of discourse semantics, in terms of 'attitude' (feelings and evaluation of things), 'engagement' (understanding how opinions are placed in discourse) and 'graduation' (the grading of feelings) (Martin and White, 2005). The appraisal system network is shown in *Figure 3-5* (see *Methodology Chapter Section 4.3* for a more detailed explanation of appraisal including further sub-systems). This system describes regions of meaning encompassing attitude, engagement, and graduation, as well as considering polarity (between positive versus negative evaluations). Meanings within the 'attitude' sub-system are also specified as either explicit meanings that are clearly evident in the evaluation in the text (inscribed attitude) or implicit meanings that require further context in order to understand (invoked attitude) (Martin and White, 2005).

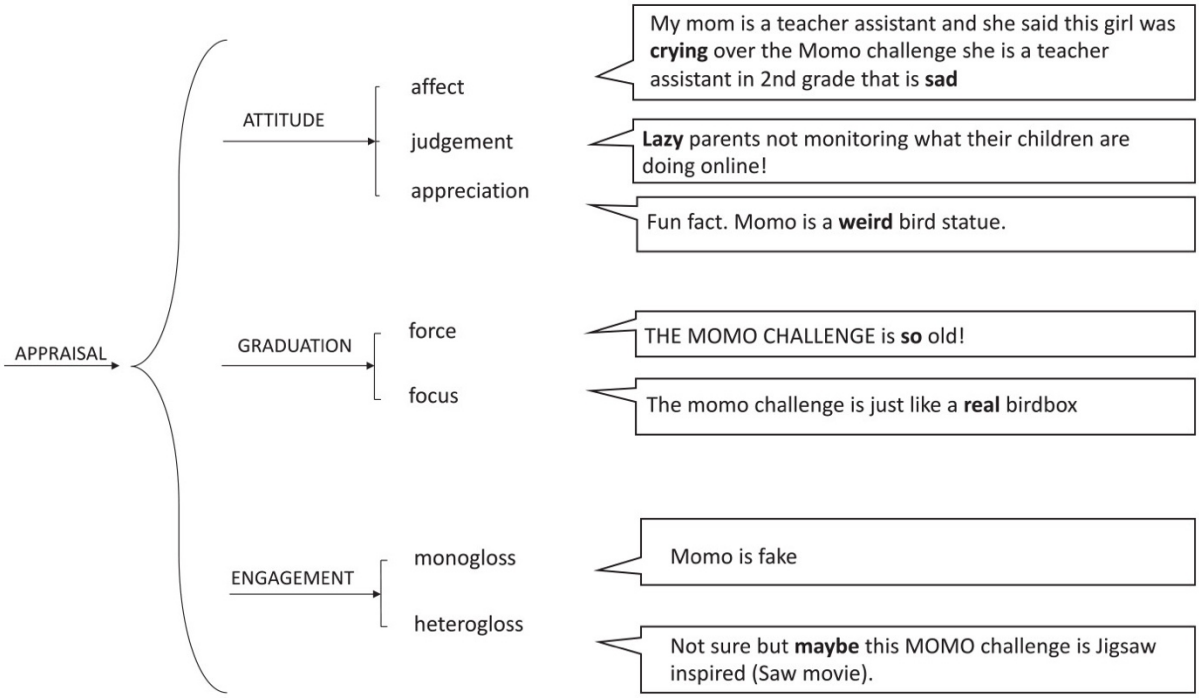


Figure 3-5 - Appraisal System Network (adapted from Martin and White, 2005) with Examples from the Thesis Dataset

In terms of metafunctions, appraisal is concerned with the interpersonal – how relationships (e.g., contact and solidarity) are enacted in discourse. By applying the appraisal framework to information disorders on social media, the interpersonal dimension of information disorders can be analysed by understanding what is at stake in the discourse (the target of evaluation) and the most common evaluations towards these ideational targets. As will be further explained in *The Methodology Chapter*, appraisal forms the basis for understanding affiliation and coupling, concepts that are central to understanding the interpersonal significance of information disorders, in terms of how people bond around certain values expressed via evaluations towards specific ideational targets.

Before the appraisal framework was developed, interpersonal meaning was an under-studied area of research. The early models of appraisal took inspiration from earlier work by Labov (1972) on interpersonal meaning in narratives and Biber and Finegan (1989) on styles of stance in English. The appraisal framework also emerged from work

on affect from the tenor register (Martin, 1992: 523-535), infant protolanguage (Painter, 1996), and work within educational research (Christie and Martin, 2009; Feez et al., 2008). Poynton (1984, 1996) contributed to research regarding affect and vocatives. Martin and White's book *The Language of Evaluation* developed the system network of appraisal that is used today (Martin and White, 2005). Their appraisal system provided a systematic account of appraisal as divided into the sub-systems of attitude, engagement, and appreciation, and further levels of delicacy. The appraisal system has been described as "the only systematic, detailed and elaborate framework of evaluative language" (Bednarek, 2006: 32). This explanation of the history of appraisal highlights how the appraisal framework has been widely used in order to study interpersonal meaning.

The appraisal framework has been applied to a range of studies, highlighting the applicability of this framework for interdisciplinary research. For example, in humorous language (Knight, 2010a), newspaper editorials (Liu and Hood, 2019), academic discourse (Hood, 2010), Youth Justice conferencing (Zappavigna and Martin, 2017), and business writing (Szenes, 2017). The appraisal framework has also been previously applied to social media research, a shared focus of this thesis, in areas such as affiliation and social tagging (Zappavigna, 2018). Zappavigna (2018) has applied the appraisal framework to tweets about fake news and how participants negotiate values around the tag #fakenews. Other work exploring values in discourse using appraisal has considered values in social media discourse (Zappavigna, 2012; Zappavigna, 2018), social tagging (Chiluwa and Ifukor, 2015), identity (Vásquez, 2014), narrative (Page, 2003) and solidarity-building (Drasovean and Tagg, 2015). Thus, appraisal is an applicable framework that forms the building blocks of understanding interpersonal meaning in discourse, the focus of the research questions in this thesis.

3.2.2. Affiliation and Bonding

The study of social alignment through the SFL concepts of affiliation and bonding is important to this thesis because social bonding is central to the spread of information disorders online. The ways in which users' bond around certain content online reveals

what is actually at stake beyond a simple view of the content of the misinformation. Ideation-attitude couplings are the unit of analysis applied in affiliation theory, as developed by Knight (2010a). As an analytical unit, a coupling refers to the "combination of meanings across a range of semiotic dimensions 'as pairs, triplets, quadruplets or any number of coordinated choices'" (Martin, 2008b: 39; Zappavigna, 2018: 205). The ideation-attitude coupling, linking attitude to an ideational target (Zappavigna, 2018: 112), gives us a way of understanding not only how people express opinion and emotion, but how they express opinion and emotion "about people, places and things, and the activities they participate in, however abstract or concrete" (Martin, 2008a: 58). This makes ideation-attitude couplings a useful unit for understanding the key ideas and attitudes that user's hold and how these are involved in affiliation strategies.

In order to understand affiliation, it is also important to understand 'bonding' as this term has a particular theoretical underpinning in SFL. Stenglin developed the notion of bonding in her research regarding three-dimensional spaces, describing bonding as how people can build "togetherness, inclusiveness and affiliation" around spatial elements such as bonding icons (Stenglin, 2004: 402). Bonding icons, also referred to as 'bondicons' refer to symbols that represent the ideologies of the people they belong to (Martin and Stenglin, 2007) and are used by communities of people to rally around. Bonding also occurs in language, particularly by users deciding to align around a particular attitude. Thus, there is a strong connection between the concepts of bonding and affiliation in terms of describing the alignment of values among communities. As Knight writes bonding refers to "the social semiotic units that bring us together" and affiliation is "a social process to account for the various ways that we construe social bonds together beyond communing" (Knight, 2010a: 71). Thus, bonding refers to a single bonding icon that a community may rally around, whilst affiliation encompasses the social process of bonding and the various ways in which this can be realised linguistically.

Knight's (2010a) research examined how people share, reject or laugh off particular bonds in humorous conversational exchanges. A main finding from this research was that affiliation discursively constructs who we are and who we aren't, thus it is an

important resource in communicating with others (Knight, 2010a). Zappavigna's concept of 'ambient affiliation' developed Knight's dialogic framework to account for online communication, where "the affiliation is ambient in the sense that the users may not have interacted directly and likely do not know each other and may not interact again." (Zappavigna, 2011: 801). The spread of misinformation online is often ambient, in the sense that users may bond around hashtags targeting particular ideas and directing them to an ambient audience. The affiliation may also be dialogic as in YouTube video comments where users are directly replying to one another.

In this thesis, there are two forms of affiliation that will be explored. The first type is dialogic affiliation that began with Knight's model of affiliation developed in work on conversational humour (Knight, 2008; Knight, 2010a; Knight, 20130b; Knight, 2013). This model identified three main categories of dialogic affiliation: rallying, laughing and condemning. Work by Zappavigna (2018) augmented the dialogic affiliation framework to encompass online communication. Zappavigna's (2018: 129) dialogic affiliation framework considers the choice "between tabling a coupling or ignoring an interlocutor entirely". When the coupling is tabled, there are further selections such as supporting, rejecting, or ignoring the bond (refer to *The Methodology Chapter* for a system network of dialogic affiliation).

The second type of affiliation is the communing affiliation framework. This framework accounts for the ambient aspect of online communication where users do not interact directly. Instead of conversation-like interaction, "ambient affiliation involves communing around, rather than necessarily directly negotiating, particular couplings in a clearly defined conversational exchange" (Zappavigna, 2018: 132). The original communing affiliation framework encompasses three main sub-systems: convoking (mustering around community), finessing (heteroglossically positioning a coupling in relation to other potential couplings) and promoting (interpersonally emphasising a coupling) (Zappavigna, 2018; Zappavigna and Martin, 2018; Zappavigna, 2021). In newer versions, promoting has been renamed tempering in order to account for how a coupling can be both emphasised or de-emphasised (Doran et al., forthcoming; Inwood and Zappavigna, 2021). Overall, the dialogic and communing affiliation frameworks

detail how affiliation strategies are realised in discourse, in terms of explaining the particular linguistic resources that may act upon a coupling, influencing how the bonds instantiated in these couplings are shared or contested.

3.2.3. Discourse Iconography and Iconisation

Discourse iconography is an aspect of SFL relevant to this thesis, because it is concerned with how communities include or exclude certain ideas and people. In other words, discourse iconography is a model of how discourse enacts communities that rally around particular 'icons' and 'values', and was developed in Tann's (2011; 2012) research into Japanese national identity. In particular, Tann developed a tripartite model to account for identity in language, consisting of 'Geminschaft' (how discourse enacts communities as fellowships that include and exclude), 'Doxa' (the communal values that are rallied around) and 'Oracle' (the iconic artefacts, images and texts that represent these values), as illustrated in Figure 3-6 (Tann, 2011). Zappavigna and Martin (2017) adopted Tann's model for their work on Youth Justice Conferencing. In this adapted version of Tann's model, 'Geminschaft' is renamed 'Communitas' in order to account more for a sense of community, which was particularly important in the context of Youth Justice Conferencing (Zappavigna and Martin, 2017). 'Oracle' is also adjusted in order to account for icons and creeds, therefore, in Zappavigna and Martin's model, 'Oracle' is renamed 'bondicon' in the tripartite model (Zappavigna and Martin, 2017). This is illustrated in Figure 3-7.

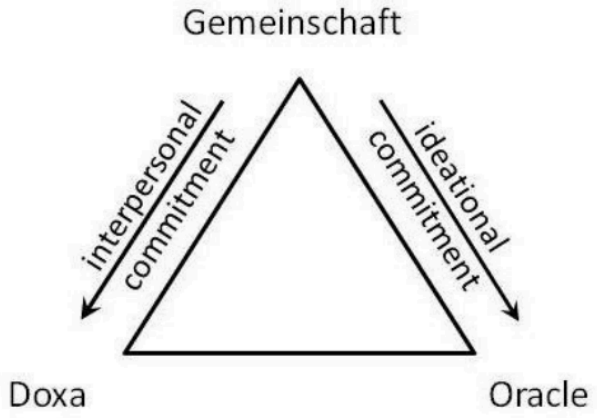


Figure 3-6 – Topology of Identity Icons from Tann (2010)

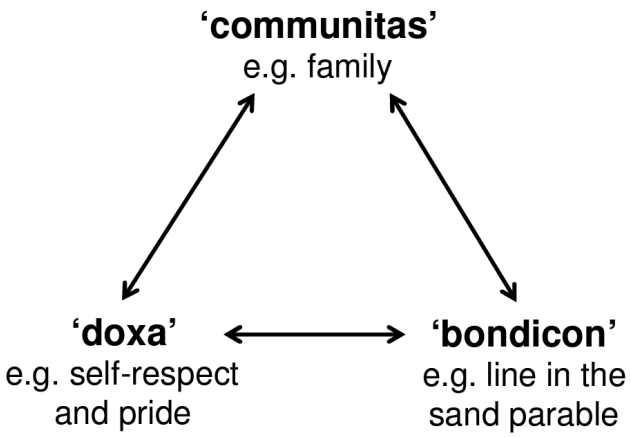


Figure 3-7 – Reinterpreted Topology of Identity Icons from Martin and Zappavigna (2013)

Iconisation is central to Tann's model and refers to "the process whereby the everyday meaning of an event or an entity is backgrounded while its emotional significance to members of a group is foregrounded (or vice versa)" (Zappavigna and Martin, 2017: 273). Iconisation involves ideational meaning being discharged whilst interpersonal meaning is charged. For example, when Trump said in his 2016 Presidential Campaign that he would 'build the wall', the ideational meaning of the wall did not take precedence, rather the interpersonal meaning of 'keeping migrants out of America' became the

predominant meaning. In this sense, an object or icon is imbued with specific interpersonal meanings that become more important than the object itself.

Iconisation was also inspired by Stenglin's work on bonding that explored how museum visitors rally around symbolic icons such as flags and memorabilia (Stenglin, 2004). In this way it is also linked to the concept of bonding icons (bondicons) that can be rallied around. To briefly illustrate, the Notre Dame Cathedral could be interpreted as a bonding icon because its interpersonal significance (an importance place of worship for Roman Catholics and recognisable symbol of France) outweighs its ideational significance (a building located in Paris). Additionally, iconisation can also involve people such as Gandhi representing peace (Zappavigna and Martin, 2017: 274). In all, discourse iconography and iconisation represent the significance of bonding icons as markers of identity and as strategies of affiliation for communities to rally around.

3.2.4. Textual Personae

The concept of a textual persona is an important notion to consider in this thesis because it provides a way of generalising the linguistic patterns observed in social media comment threads. This is particularly helpful when wanting to understand how people engage with information disorders. Textual personae from an SFL perspective, refers to "persons and personalities communing in discourse" (Martin, 2009). However, a persona does not refer to an individual person, but rather is a generalisation of how identities are enacted linguistically. These identities are characterised by particular "coupling dispositions" (Zappavigna, 2014a; Zappavigna, 2014b; Zappavigna, 2018), that is, a tendency to construe particular patterns of values. The linguist Firth (1950) originally referred to the idea of "bundles of personae", that is, how personalities interact in discourse and how one person can enact multiple personae depending on the circumstance they are in. A similar position is adopted by Don (2018: 72) in their definition of textual personae as "a function of the acts of positioning that a writer makes, or, in mediated interactive contexts the positioning each contributor makes to a discussion in response to others' comments". In this way, a textual persona is

distinguished by these positioning acts rather than any individual's self-image or social role ascriptions (Don, 2018: 72).

This thesis argues that by considering personae in terms of their tendency to negotiate particular values, we can understand the main ideologies that personalities bond around in the context of misinformation. Work considering the function of textual personae in SFL has focused on the construal of personae at different levels of linguistic stratification. For instance, discursive personae have been examined from the perspective of dialogistic positioning, drawing on the engagement sub-system of appraisal (White and Sano, 2006), authorial personae have been examined through evaluative disposition (White, 2008), textual personae have been explored at the level of discourse semantics (Don, 2007; Don, 2016) and tenor (Don, 2009; Don, 2018), and identity has been explored from the perspective of iconography (Tann, 2011; Tann, 2012). SFL explorations of personae have also examined shared values in microblogging (Zappavigna, 2014a), as an act of impersonation using paralinguistic and dialogic resources (Logi and Zappavigna, 2021), and as communal identity in conversational exchanges (Knight, 2010b). The work in this thesis interprets personae from the discourse semantic level of SFL, focused on interpersonal meaning enacted by resources from the appraisal and affiliation systems.

Beyond studies in SFL, personae have been explored in other linguistic fields, and in media and cultural studies. In linguistics outside of SFL, researchers have used the term 'persona' and have explored it according to phonology (D'Onofrio, 2018), for example dialect style as a type of persona management (Coupland, 2002) and as a register of language from a linguistic anthropology perspective (Agha, 2005). Goffman's work on the presentation of the self and impression management also aligns with the notion of a persona (Goffman, 1978). In media studies, personae have been explored in terms of how one can enact different identities online and offline (du Preez and Lombard, 2014) and from the perspective of 'advertising persona', defined as a created character "whose purpose is to persuade an audience to consume" (Stern, 1994).

The specific field known as 'persona studies' takes a cultural studies perspective and connects the study of persona with the "close study of the performance and assemblage of the individual public self" (Marshall et al., 2019: 17). In the move from researching persona to developing the field of persona studies, Marshall, Moore and Barbour write that the aim of persona studies is "experimenting with and developing a set of approaches for analysing the expansion and proliferation of the development of the public self" (Marshall et al., 2015: 289). The authors acknowledge that this recent interest in studying persona has emerged from studies of social media and micro-celebrity (Senft, 2013), positing that since an individual can present their identity online in multiple ways, the study of persona is particularly important.

In comparison to the conventions used in this thesis, Marshall, Moore and Barbour refer to 'personas' rather than 'personae'. From their perspective, 'personas' refers "to the multiple aspects of an individual's character that are presented and understood by others at certain times and places and in certain roles; a politician, a mother, a celebrity" (Marshall et al., 2015: 302). They interpret 'personae' as "reserved by dictionaries" and used to refer to "members of a dramatic work or the multiple characters inhabited by the author over a series of novels" (Marshall et al., 2015: 302). As the research surveyed shows, there are some differences in terminology across different disciplines and personae have been explored from a diverse range of perspectives. Whilst this thesis takes inspiration from these various disciplines, it maintains a social semiotic conception of personae which coordinates with the SFL analytical methods applied to the dataset.

3.2.5. Instantiation and Individuation

In order to further understand personae from an SFL perspective, two particular hierarchies within SFL, instantiation and individuation, are relevant. The cline of instantiation refers to the relationship between a language system and an instance of text. In other words, instantiation is a 'scale of generalisation' with system at one end (generalised meaning potential) and reading at the other end (subjectified meaning) (Halliday and Matthiessen, 1999; Matthiessen et al., 2010). The cline of instantiation is

illustrated in *Figure 3-8*. Halliday and Matthiessen (1999) use the analogy of weather and climate to explain instantiation: weather and climate are the same phenomenon but looked at from different perspectives. Climate is the generalisation of weather patterns whilst weather is an instance of climatic trends (Halliday and Matthiessen, 1999). Thus, instantiation shows us how the same meaning can be seen from different perspectives in terms of patterns of generalisation.

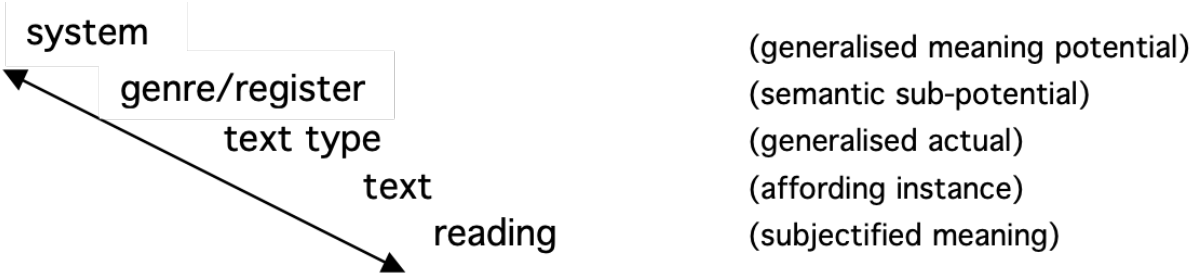


Figure 3-8 - Cline of Instantiation (adapted from Halliday and Matthiessen, 1999)

Alongside instantiation, is individuation which refers to the "relationship between the reservoir of meanings in a culture and the repertoire a given individual can mobilise" (Matthiessen et al., 2010: 117). The individuation hierarchy is shown in *Figure 3-9*. Individuation takes inspiration from Bernstein's work on coding orientation (Bernstein, 2003). This concept has been used in SFL for exploring semantic variation (styles) as "systematic sociosemantic variation" where the concept of meaning is "arbitrarily constrained" rather than "meaning preserving" (Hasan, 1989: 221). Individuation is organised as a cline extending from the collective to individual when analysing a text. A concept central to this thesis, persona, is realised as the personal repertoires of language users based on a pattern of instantiations. In this sense, individuation works as a complementary hierarchy to realisation and instantiation.

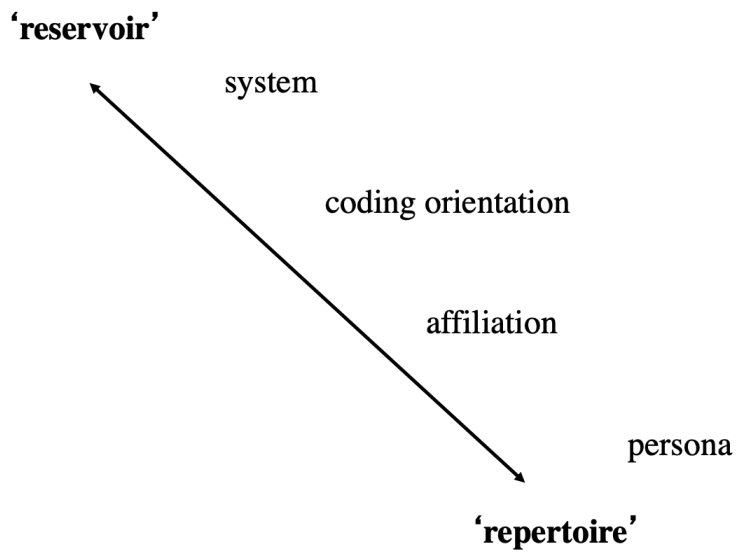


Figure 3-9 - Adapted from Martin (2006, p. 294)

Individuation is important to the study of affiliation because it provides a "perspective on identity that is informed by notions of ideology" (Knight, 2010a: 53). The individuation hierarchy in relation to affiliation is shown in *Figure 3-10*. As we can see, 'persona' is located at the bottom end of the hierarchy (micro-perspective) whilst culture is at the top end of the hierarchy (macro-perspective). The affiliation perspective on individuation shows "how resources of the culture can be deployed to members to commune, rather than allocated to individual repertoires" (Knight, 2010a: 57). In Knight's work on the cline of individuation, bonds are at the level of 'persona', bond networks at the level of 'sub-culture', ideological networks at the level of 'master identity' and culture is a 'system of bonds' (Knight, 2010a: 238). Overall, this thesis considers textual personae at the level of repertoire based on the cline of individuation, and in terms of bonds based on Knight's work.

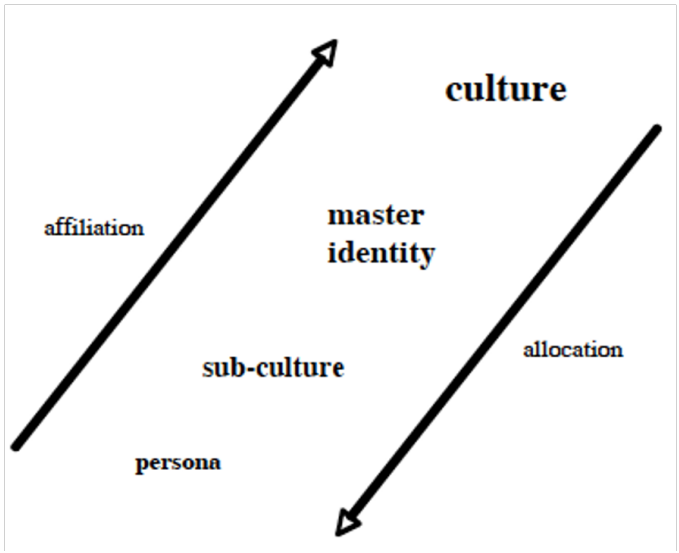


Figure 3-10 - Adapted from Knight (2010, p. 57)

3.2.6. Genre

As indicated previously in the literature review chapter, deceptive communication and information disorder function as umbrella terms, encompassing the many different ways in which the truth can be distorted. In this sense, genre is another concept that is relevant to this research, in terms of mapping out these different forms of deceptive communication and information disorder. According to the SFL-based genre framework, deceptive communication and information disorder are best approached as macro-genres, since they involve configurations of elemental genres. This section will provide an overview of genre in SFL terms, particularly focused on Martin and Rose's work.

In SFL terms, genre can be understood as how society shapes language. Halliday and Hasan (1989) considered genre as part of register, and primarily used the term to refer to literary genres. Martin and Rose's work on genre extended SFL's hierarchy of strata by further stratifying context into genre and register (Martin and Rose, 2008). According to this perspective, genres are "a recurrent configuration of meanings" and we need "to think about more than individual genres; we need to consider how they relate to one another" (Martin and Rose, 2008: 5). Genres are characterised as "staged,

goal oriented social processes": staged because they consist of multiple steps, goal orientated because they are created in order to fulfil a purpose, and social because they are created for readers/audience in a particular social context (Martin and Rose, 2008: 5).

Martin and Rose's work on genre developed from their 'Sydney School' research into the literacy needs of primary school students (Martin and Rose, 2008). Thus, the genres discussed in Martin and Rose (2008) directly relate to genres that primary school children are taught to write. This consists of story genres such as recount, exemplum, and narrative, and reports such as explanations and scientific genres. Within these identified genres, there are particular stages. For example, a narrative consists of a complication, evaluation and then a resolution. In addition, 'macro-genres' are defined by Martin and Rose (2008: 201) as "short genres that go to make up larger texts". In this sense, a macro-genre could consist of both narrative and exemplum stages. Therefore, genres and macro-genres can be realised as multi-layered and distinguished by specific linguistic configurations.

Martin and Rose's work on genre has inspired a range of research in SFL and beyond. Hao and Humphrey (2009) have investigated the co-patterning of appreciation and ideation in published research warrants, Christie and Derewianka (2010) have researched school discourse across multiple years of schooling, Flowerdew and Wan (2010) have researched the company audit report genre, Maton et al. (2015) have researched genre in relation to legitimation code theory, and Dreyfus et al. (2015) have researched genre pedagogy in higher education. Genre has also been explored from the perspective of multimodality studies, particularly in work by Bateman on genre and multimodality, encompassing computational approaches to analysing the moving image (Bateman, 2013) and multimodal documents (Bateman, 2008). Recent research has also emerged in relation to analysing YouTube promotional videos with Martin and Rose's (2008) genre approach combined with Kress and van Leeuwen's (2006) multimodal discourse analysis (Moreno López, 2020; Qisthi et al., 2022; Suriyanti, 2022). These approaches to genre in multimodality studies will be discussed in more detail in the following section.

3.3. Multimodal Discourse Analysis

Multimodal Discourse Analysis (MDA) draws on the strategies for analysing linguistic meaning provided by SFL, to consider how other semiotic modes can be modelled such as visual images and other kinds of paralanguage (e.g. gesture). Gunther Kress and Theo van Leeuwen are leading figures in multimodality studies. Their work on developing a systemic grammar of visual images is encapsulated as *Reading Images: The Grammar of Visual Design* (Kress and Van Leeuwen, 2006). An MDA approach has previously been applied to visual images (Caple, 2008; Kress and Van Leeuwen, 2006; O'Toole, 1994), sound and music (Caldwell, 2010; Van Leeuwen, 1991; Van Leeuwen, 1999), architecture (Kress and Van Leeuwen, 2006; O'Toole, 2004; Ravelli and McMurtrie, 2015; Ravelli, 2000; Stenglin, 2004), gesture (Hood and Lander, 2016; Martinec, 2004), and facial expression (Tian, 2010). MDA research which adopts and SFL approach considers semiotic resources as simultaneously enacting the three metafunctions explained earlier in this chapter. This section will now detail how these metafunctions have been recontextualised to account for a multimodal perspective on meaning that factors in resources beyond language, of the kind that are important to understanding the YouTube videos in this thesis.

Multimodal research methods focused on the ideational metafunction have provided a way of understanding how experience is construed, which is useful for researching environments that contain instances of information disorder. The ideational metafunction refers to "the world around and inside us" (Kress and Van Leeuwen, 2006: 55). For example, one aspect of the ideational metafunction used to explore semiotic resources multimodally is narrative representation. This is concerned with structures which represent "aspects of reality in terms of unfolding actions and events, processes of change, transitory spatial arrangements and so on" (Kress and Van Leeuwen, 2006: 55). These sorts of narrative structures contain a vector that realizes a dynamic process and consider the relations between participants (Kress and Van Leeuwen, 2006: 55). The ideational metafunction also focuses upon conceptual representations, which refers to "representing participants in terms of their more generalized and more or less stable and timeless essences" (Kress and Van Leeuwen, 2006: 76). Conceptual representations

are focused on ideas such as symbolic processes that consider how a participant's identity is established (Kress and Van Leeuwen, 2006: 102). The ideational metafunction shows how multimodal resources can be analysed in order to interpret their symbolic meaning and how they construe experience.

Multimodal research methods have also focused on the interpersonal metafunction as a way of understanding communication strategies and credibility, concepts that are particularly important to consider when studying information disorders. The interpersonal metafunction is concerned with how semiotic resources enact relationships. Multimodally, as Kress and van Leeuwen write "images involve two kinds of participants, represented participants (the people, places and things they depict) and interactive participants (the people who communicate with each other through images, the producers and viewers of image)" (Kress and Van Leeuwen, 2006: 113). Thus, the interpersonal metafunction provides a way for understanding how gaze and different angles construe power. Another area of interest regarding the interpersonal metafunction in multimodal research is the concept of modality or validity. Modality is concerned with the "signs of credibility" displayed in visual communication (Kress and Van Leeuwen, 2006: 149). Kress and van Leeuwen state that modality represents the interpersonal metafunction because "it does not express absolute truths or falsehoods: it produces shared truths, aligning readers or listeners with some statements and distancing them from others" (Kress and Van Leeuwen, 2006: 151). Modality has previously been considered by Kress and Van Leeuwen (2006) in images, in relation to the coding orientations of 'scientific/technological', 'abstract', 'naturalistic' or 'sensory'. These coding orientations can provide a greater understanding of how images are truthful or deceitful in relation to the genre they belong to e.g. a scientific/technological coding orientation would need to have an absence of background and minimal colour, more comparable to a blueprint than a naturalistic image, to be of the highest modality in terms of being an acceptable scientific image. In the third edition of *Reading Images*, Kress and van Leeuwen have renamed modality 'validity' in order to avoid confusion between mode as the "means/technology of representation" and mode as "the meaning of assigning something like truth the value of an utterance" in Hallidayan grammar (Kress and Van Leeuwen, 2006: xvii). Additionally, since the first publication of *Reading*

Images, more consideration has been drawn to the role of modality in digital communication. As Ravelli and Van Leeuwen (2018: 284) have remarked "the digital age is subtly changing the look of everyday photographic images" and is altering our notions of photographic realism, thus keeping this in mind, new validity standards will continue to be created, particularly as the sensory and abstract coding orientations continue to combine (Kress and Van Leeuwen, 2006: 178). In all, there are many different multimodal methods that can be used to explore the interpersonal metafunction, more of these methods will be explored in greater detail in the following sections.

Lastly, research methods that encompass the textual metafunction are particularly useful when studying online media such as YouTube, that consist of multi-layered structures and elements. The textual metafunction is concerned with how language and other semiotic modes are organized. In Kress and van Leeuwen's terms, the textual metafunction is associated with composition and "relates the representational and interactive meaning of the image to each other through three interrelated systems" (Kress and Van Leeuwen, 2006: 181). These systems consist of information value (the specific values that are associated with the placement of elements), framing (how framing devices form meaning) and salience (what elements attract the viewer's attention (Kress and Van Leeuwen, 2006: 182). The textual metafunction has been explored particularly in regards to research on multimodal documents (Bateman, 2008) and websites (Djonov, 2005), and will be further discussed in Section 3.3.

Overall, the metafunctional approach to multimodality, led by Kress and van Leeuwen's work, has provided a rich array of methods to study multimodal phenomena. In the following sections, particular aspects of multimodality that are relevant to the conceptual aims of this thesis will be explored. This includes studies that have applied van Leeuwen's legitimation framework (Van Leeuwen, 2007), relevant issues regarding the exploration of identity from a multimodal perspective (Van Leeuwen, 2021), multimodality and the moving image particularly focused on Bateman's and O'Halloran's work (Bateman, 2008; Jewitt et al., 2016), the discursive analysis of news values created by Bednarek and Caple (2017), and the intersections between multimodality studies and critical discourse analysis as explored by Van Leeuwen

(2008). Whilst this does not cover every element of multimodality studies, these sections identify the most important literature to consider regarding the aims of this thesis, to study information disorders from a multimodal perspective and a perspective that takes into account the importance of identity when applying a social semiotic approach.

3.3.1. Legitimation

Legitimation is an important concept underpinning this thesis, as the strategies people use to legitimate or delegitimate information can show how deceptive or problematic content can present itself to be credible. There is already a history of legitimation research in discourse studies and particularly multimodal discourse studies. In discourse studies, Van Dijk (1998) refers to legitimation in terms of 'institutional contexts', highlighting the interaction between legitimation and legitimacy. In Van Dijk's work, it is institutions that create legitimacy and regulate what is deemed legitimate. Van Leeuwen's approach to legitimation is different, in the sense that legitimation refers to how discourses construct credibility, in other words "why should I do this?" or "why should we do this in this way?" (Van Leeuwen, 2007: 94). Thus, van Leeuwen's approach to legitimation exists beyond just institutions. Instead, there are four main categories of legitimation: authorisation (legitimation in relation to an authority figure or tradition), moral evaluation (legitimation in relation to value systems), rationalisation (legitimation by reference to social actions and knowledge constructed by society) and mythopoesis (legitimation expressed by narratives that reward legitimate actions) (Van Leeuwen, 2007: 92). Whilst van Leeuwen's earlier work on legitimation existed at the lexico-grammatical level (Van Leeuwen, 1995), the most well-known legitimation framework by Van Leeuwen (2007: 92) considers legitimation as realised by "specific linguistic resources and configurations of linguistic resources". Although some of these resources can be lexico-grammatical, legitimation is also realised at the discourse semantic level and multimodally. This flexibility of the legitimation framework to be applied at various strata and modes, has meant that it has been a widely applied framework.

The legitimation framework has been applied across a range of different studies with different methodological adjustments. Multimodal legitimation has been applied to campaigns and political parties (Chaidas, 2018; Mackay, 2015), school history books (Peled-Elhanan, 2010) and news reports (Pérez-Arredondo and Cárdenas-Neira, 2019). Legitimation on a rhetorical level has been explored via the use of metaphors and frames (Hart, 2017; Näsänen, 2017), narratives (Chaidas, 2018; Näsänen, 2017), and via a lexico-grammatical analysis (Oddo, 2011). More hybrid discourse approaches to legitimation have also been applied, for example, to studies focused on discourse-historical analysis (Van Leeuwen and Wodak, 1999) and New Rhetoric (Erkama and Vaara, 2010). In terms of social media studies and discourse, legitimation has been applied to political fake news in Nigeria (Igwebuike and Chimuanya, 2020), Pakistani social media discussions (Rizwan, 2019), and the use of Facebook by an English language classroom (Vanek et al., 2018). Delegitimation (when linguistic strategies are used in reverse to discredit) has been considered in relation to political memes from a multimodal perspective (Ross and Rivers, 2017), tweets from a discourse perspective (Rivers and Ross, 2020; Ross, 2020), and discourses about the delegitimation of migration (Sundström and Obenius, 2020). These examples demonstrate how adaptable the legitimation framework is and the multiple ways in which it has been adjusted to meet specific research aims. As will be further explained in *The Methodology Chapter* this thesis has also adapted the legitimation framework to account for the multimodal nature of YouTube data.

3.3.2. Multimodality and Identity

Issues of self-representation and identity have been a focus in MDA studies and are issues that this thesis also considers when investigating how YouTuber's represent themselves and their ideologies, and persuade others to believe their claims via a range of multimodal resources. Van Leeuwen (2021: 6) considers four types of semiotically realised identity: social identity (identity in relation to our place in the social order), individual identity (identity as part of our inner world), role identity (having as many identities as the roles we hold in life) and lifestyle identity (an identity that has emerged in contemporary life and is related to a focus on leisure and consumer life). Van Leeuwen argues that these types of identity have different semiotic realisations (Van

Leeuwen, 2021: 7). For example, role identity is realised by "emotive investment" (acting with feeling), "social recognition" (socially ratified identity), and "multiplicity" (changing from one identity to another) (Van Leeuwen, 2021: 18). Van Leeuwen also uses the term 'role styles' to refer to "ways of performing roles, usually in more or less institutionalised settings" (Van Leeuwen, 2021: 23). Van Leeuwen's work on identity, specifically role identity, shows how identity is semiotically and multimodally realised, and that we hold different semiotically realised identities depending on the situation we are in.

We can also think more specifically about how identity on social media has been researched from an MDA perspective, as this thesis is focused on analysing the visuality of YouTube videos from the perspective of identity and ideology. Zhao and Zappavigna (Zappavigna and Zhao, 2017; Zhao and Zappavigna, 2018a; Zhao and Zappavigna, 2018b) have conducted research on selfies from an MDA perspective in order to understand the interpersonal significance of selfies. As they write their approach foregrounds the intersubjective function of the selfie, where the focus is not just "this is me" but rather "this is my perspective" (Zhao and Zappavigna, 2018a), thus a selfie holds more multimodal significance than just an act of vanity. Sindoni (2020) has explored how identities are entextualised from a systemic-functional and multimodal discourse analysis perspective by health providers and peer carers when using digital platforms that focus on mental health issues. Koteyko and Hunt (2016) have studied how health identities are performed on social media, specifically looking at the multimodal resources used when interacting on Facebook. In all, these examples show how an MDA approach is useful for analysing identity on social media, that aligns with this thesis' aims to analyse textual personae on YouTube.

3.3.3. Multimodality and Computational Studies

MDA studies have been particularly focused on applying computational methods to understanding images from a semiotic perspective and these research concerns are also relevant to this thesis because constructing large datasets that can interpret information disorders requires approaches that align with computational focuses. Bateman (2008:

15) developed the GeM Model (Genre and Multimodality Model) for analysing multimodal documents at a large scale and to define “several layers of description for multimodal documents” and has expanded this to analysing moving images. This is an empirical approach, that involves understanding the social practices and affordances of documents. Bateman's GeM Model considers three key textual dimensions: "text flow" (how text moves within the videos), "image flow" (how images move within the videos) and "page flow" (how the frames within the videos move in succession) (Bateman, 2008: 175). A diagram of these three semiotic modes is shown in Figure *3-11*. Understanding information disorders as a multimodal genre provides a greater understanding of the complexities of false information and issues such as structure, ideology, and how different texts interact with one another.

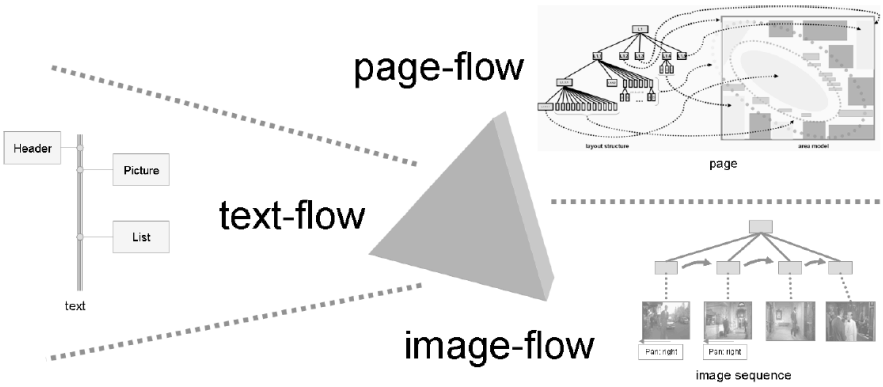


Figure 3-11 – Three Semiotic Modes Deployed Within Document Pages (from Bateman, 2008, p.175)

MDA and computational approaches have also been applied to social media, which is a central concern of this thesis. Bateman et al. (2017) consider how social media platforms can be analysed via a computational MDA approach. Using Instagram as an example, they consider the platform as a ‘canvas’ that can be manipulated (Bateman et al., 2017: 362). This canvas is then divided into further subcanvases, separating the interface of the platform (from the consumer’s perspective) from the programmatic semiotic modes of the platform (from the producer’s perspective). Subcanvases also exist for the consumer who can see both the content canvas and commentary canvas of the platform.

Thus, by considering the 'canvas' this allows us to analyze the different semiotic modes that can exist on a single platform, and can be applied to qualitative and quantitative analyses of social media. However, Bateman et al. (2017) also point out that due to social media data being "inherently noisy" with content such as screeenshots and memes this can pose additional challenges to quantitative analyses that rely on automatic analysis techniques (Bateman et al., 2017: 365).

Computational approaches to multimodality have also been concerned with how the principles underpinning SFL can be automated, which is of interest to this thesis, as algorithmic approaches to detecting misinformation have frequently been discussed (refer to *The Literature Review Chapter* for an overview of these different approaches and concerns). Kay O'Halloran has contributed to Systemic Functional-Multimodal Discourse Analysis (SF-MDA) studies by developing systems to analyse visual images through a metafunctional lens (O'Halloran, 2008), combining qualitative methods of MDA with quantitative methods of data mining and information visualisation (O'Halloran et al., 2018), and the visual semiosis of film (O'Halloran, 2004). O'Halloran has also applied mixed methods approaches to analysing the text and image relations in violent extremist discourse (O'Halloran et al., 2019; Wignell et al., 2021; Wignell et al., 2017a) and changes over time in extremist magazines (Wignell et al., 2017b). In all, computational approaches to multimodality have provided new methods for particularly understanding images from the textual metafunction, and how creating a combined qualitative and quantitative analysis is a useful approach to analysing large datasets of images and videos.

3.3.4. Multimodal Discourse Approaches to Analysing Social Values

This section focuses on a final concern of this thesis, MDA approaches to analysing social values, in other words, research methods that have focused on understanding how values are construed in society. MDA approaches to analysing social values are another area of interest for this thesis, because social values are used to affiliate with others and can help perpetuate information disorders. This section will focus on two

main approaches to studying social values: Discursive News Values Analysis (DNVA), and intersections between MDA and Critical Discourse Analysis.

In news media research, Bednarek and Caple (2017) created the first detailed methodology that addresses the multimodal construction of newsworthiness in terms of verbiage and images. Discursive News Values Analysis (DNVA) refers to "how news values are constructed through discourse (i.e. semiotic resources in use)" (Bednarek and Caple, 2017: 49). DNVA assumes:

> '…that material events are endowed with newsworthiness by the media, for example, by emphasising or de-emphasising certain news values in texts (Bednarek and Caple, 2014: 139). We also assume that the potential news value of events depends on a given sociocultural system that assigns them value." (Bednarek and Caple, 2017: 51).

Thus, DNVA is a semiotic approach that foregrounds the particular values associated with a text. Regarding 'newsworthiness', it consists of the following values summarised in Table 3-1. These news values (excluding aesthetics) can be applied to verbiage and images. Aesthetic appeal only applies to visuals dealing with how balance is constructed through "both composition and technical qualities" (Bednarek and Caple, 2017: 66). Bednarek and Caple argue that aesthetic appeal is only considered visually and not in language, because aesthetic devices are not valued in serious journalistic writing which instead values facts and clear language and visuals over aesthetic choices like overly florid language and visuals (Bednarek and Caple, 2017: 67). DNVA is a useful methodology for understanding news discourse from a multimodal perspective, this approach allows the researcher to apply a consistent framework to both visuals and language.

News Value	Brief description
Negativity/positivity	Conflict/solidarity
Impact	Consequence, significance, relevance
Superlativeness	Size, scale, scope
Proximity	Geographical, cultural nearness
Timeliness	Recency, currency
Unexpectedness	Unusuality
Eliteness	Prominence, elite status
Personalisation	Ordinariness, personal
Consonance	Expectedness, typicality
Aesthetics	Visuals only, aesthetically pleasing

Table 3-1 – News Values as Summarised in (Bednarek and Caple, 2017: 53)

DNVA has been applied by a number of researchers working in linguistics and semiotics. Makki (2019, 2020) explores the discourse of news values in Iranian crime news reporting, Huan (2016) applies DNVA to Chinese and Australian hard news stories about risk events, Dahl and Fløttum (2017) research newsworthiness in relation to stories about climate change in British newspapers and Molek-Kozakowska (2017) takes a DNVA approach to understanding popular science journalism. Overall, this is a very applicable framework that is particularly relevant to this thesis in terms of how information disorders can share many of these news values visually and verbally.

In work by Van Leeuwen (2008) multimodality has also been explored in conjunction with work in Critical Discourse Analysis (CDA), drawing particular attention to issues of power and society. This work was inspired by Michel Foucault's concept of discourses as serving the interests of particular historical or social contexts (Foucault, 2021), Bernstein's concept of recontextualization (Bernstein, 2003), Martin's theory of activity sequences (Martin, 1989; Martin, 1992), and Michael Halliday's concept of 'register' that focuses on the context of situation (Halliday, 1995). This work is particularly important for combining critical discourse analysis and multimodal semiotics to show how social actors and practices are realised.

By 'social practices' Van Leeuwen refers to "socially regulated ways of doing things" (Van Leeuwen, 2008: 6), however this 'regulation' is not just limited to institutions but can apply to many everyday aspects of life. By 'social actors' van Leeuwen refers to the "participants of social practices" (Van Leeuwen, 2008: 23). Visually, social actors can be realised by social distance (how close or far away from other people), social relations (from what angle do we see the person), social interaction (how do we describe their gaze), exclusion (by people or groups), roles (are they involved in the action or not), specific and generic (how are people represented), individuals and groups (how are people represented) and categorisation (are they cultural or biological) (Van Leeuwen, 2008: 136-148). Social actor theory has also been used in order to understand issues of visual racism in Van Leeuwen (2008) and Richardson and Wodak (2009). This approach provides a way of describing how visually communicated racism is construed and understanding how images can be used for deceptive purposes such as to include or exclude certain groups of people.

3.4. Conclusion

This chapter has provided an overview of the main theoretical concepts in Systemic Functional Linguistics (SFL) and Multimodal Discourse Analysis (MDA) that are relevant to the overarching concern of this thesis with providing an interpersonal perspective on information disorders on social media. The chapter has traced the early history and key concepts underpinning SFL. SFL concepts that are particularly pertinent to the analysis undertaken thesis were detailed, spanning appraisal, affiliation, discourse iconography, textual personae, and genre. These concepts illustrate the focus of this thesis on using the tools of discourse semantics and the perspective of individuation in order to understand issues of information disorders on social media.

The second part of the chapter focussed on MDA. The concepts of legitimation, identity, computational approaches to studying images, discursive news values analysis, and critical discourse analysis were discussed. These concepts are also crucial to the focus of this thesis, particularly in regards to analysing the credibility of multimodal content

on YouTube and the strategies used to construct a sense of legitimacy and newsworthiness. Overall, this chapter shows the rich history of SFL and MDA concepts that this thesis will draw on.

4. CHAPTER FOUR DATA AND METHODOLOGY: ANALYSING YOUTUBE VIDEOS AND COMMENTS

4.1. Introduction

This chapter details the methodology underpinning this thesis and is divided into two key sections: a description of how the data was collected and sampled, and an explanation of how the data was analysed in terms of affiliation and legitimation. The aim of this thesis is to contribute to the body of research on information disorders on social media by using discourse analysis methods developed in Systemic Functional Linguistics to analyse the interpersonal meaning construed in YouTube videos and comments within two case studies.

This chapter explains the two types of discourse analysis that were applied to the dataset: an affiliation analysis for describing the strategies through which social bonds are negotiated and a legitimation analysis for exploring how these social bonds are positioned as credible. This was achieved by firstly annotating YouTube videos and comments according to the appraisal framework in terms of attitudes expressed, the multiple voices involved in discourse, and how graduation is used as a resource when conveying information. As will be further explained in this chapter, the affiliation and legitimation analysis was constructed from these building blocks of appraisal.

4.1.1. Methodological Motivations

This section will briefly outline why particular methodological decisions were made and why YouTube was selected as a data source for this thesis. As mentioned in *The Literature Review* chapter, YouTube has been an understudied social media platform in relation to information disorders compared to platforms like Twitter. In considering YouTube methodologies "starting from the same data, every study that is using YouTube as its main data source can plan a brand new research strategy combining

different metrics" (Giglietto et al., 2012: 148) due to the complex nature of the platform, including a wide range of data sources, diverse types of user engagement, multi-device usage, and the influence of recommendation algorithms. Thus, there is a lot of scope to analyse YouTube multimodally and adapt methodological frameworks for analysing YouTube.

A multimodal affiliation and legitimation analysis was applied for this thesis in order to account for the multimodal nature of YouTube data and provide a holistic understanding of the visual and linguistic strategies used by YouTubers engaging with information disorders. The analysis of both YouTube videos and YouTube comments allows for an understanding of how YouTube creators propagate social bonds as well as how YouTube commenters respond to these social bonds. Additionally, conducting both an affiliation and legitimation analysis allows for an understanding of how YouTubers share social bonds, with the extra layer of the legitimation analysis, delving deeper into why these social bonds are so persuasive. The combination of these different data types and analytical frameworks addresses the research aims of this thesis as will be introduced in the next section.

4.1.2. Research Questions

The research aims of this thesis span four key areas: evaluative meaning, affiliation, legitimation and multimodality. In particular it aims to:

1. Understand the key affiliation and legitimation strategies used in the language and visual communication of YouTubers engaging with information disorders
2. Identify and analyse the key social bonds that textual personae share or contest in comments when they respond to YouTube videos that engage with information disorders
3. Compare and contrast the textual personae visible in non-politically motivated information disorders (e.g. internet hoaxes) to politically

motivated information disorders (e.g. politically-triggered conspiracy theories)

In order to address the aims, the thesis draws on discourse analysis methods from systemic functional linguistics to analyse the interpersonal dimension of information disorders, for example, how users' bond around deceptive news stories. It aims to move beyond the 'fake news binary' that only considers information as true or false (McDougall, 2019) to provide a method for understanding the spread and impact of information disorders as a set of more complicated practices, and its relation to the role of moral panics and social bonding.

In order to address these key aims the thesis addresses the following research questions (as firstly explored in the *Introduction Chapter*):

1. How can the interpersonal dimension of information disorders spread on YouTube be illuminated via a linguistic and multimodal analysis of patterns of affiliation and legitimation?
2. How are social bonds legitimated in the transcripts and visual content of YouTube videos that spread information disorders?
3. What different textual personae emerge in the comment sections of YouTube videos that spread information disorders? Can these textual personae be grouped according to social bonds and legitimation strategies?
4. What similarities and differences emerge when comparing textual personae in non-politically motivated information disorders versus politically motivated information disorders?

4.1.3. Ethical Considerations

This section will survey the ethical considerations made in relation to this thesis. Currently, there is minimal peer-reviewed work discussing the ethics of using YouTube data. As this is a relatively new area of research, strict guidelines have not yet been developed. However, some organisations such as the Association of Internet

Researchers have created guidelines for researchers (Franzke et al., 2020). In terms of previous research into YouTube, researchers in the field of medical science have not sought ethics approval for their work despite using YouTube video data to analyse topics such as cataract surgery videos (Bae and Baxter, 2018), non-suicidal self-injury videos (Lewis et al., 2012) and multivitamins (Basch et al., 2016). As Patterson (2018) summarises, there is a continuum of positions on this issue:

> "Those who share their videos on YouTube do so with knowledge that the resulting uploads are available to anyone with Internet access, and YouTube users have access to settings that will render their videos unsearchable by unknown others if they so desire. On the other extreme of the continuum, some researchers transparently grapple with the unsettling uncertainties that exist when working with social media data." (p.760).

Analysing deceptive communication present some inherent ethical dilemmas, due to the nature of this content. Fuchs (2018) writes about issues regarding internet ethics and informed consent that are relevant to this project; in terms of analysing deceptive communication it is impossible to obtain the consent of the participants, rather the greater social good of the research project outweighs the impracticality of obtaining consent from people on social media who have engaged in controversial practices such as spreading disinformation and hate speech, and would be aware that their data was made publicly available. This thesis applies a similar philosophy to Fuchs in terms of ethical considerations. All the data from this thesis has a significant level of popularity on YouTube, so it can be considered as very public and already well known, meaning that this research is not contributing to the unnecessary amplification of deceptive content.

In this thesis, identifying information about users have been anonymised in order to ensure research integrity. The text or any facial recognition in the images were anonymised in order to ensure that the YouTuber and the specific video they are associated with were not identified. Only YouTube videos that were public and received significant engagement (more than 1000 views) were selected in order to make sure that false information was not being unnecessarily amplified and that the issues of

information disorder were already in the public sphere. The phenomenon of journalists accidentally amplifying false information into the public when reporting on issues of misinformation has been analysed by the research agency First Draft[3] (Wardle and Derakhshan, 2017), and this thesis has also taken into consideration this research and the best ethical practices for ensuring false information is not unnecessarily amplified. Regarding copyright, according to YouTube's fair use policy[4], YouTube videos can be used for research purposes as this constitutes 'fair use'. I used the YouTube authorised API (Application Program Interface) to collect data, so this also means that I am not breaching any of YouTube's data collection policies[5].

4.2. Collection of Data

The dataset explored in this thesis is comprised of two case studies: The Momo Challenge and The Notre Dame Fire. It includes YouTube videos, transcripts of these videos, and comments. The particular case studies were chosen because they each represent two different aspects of information disorders: internet hoaxes and political disinformation and conspiracies spread during breaking news stories. Additionally, these case studies can be viewed through micro, meso, and macro perspectives (Jepperson and Meyer, 2011) due the varying types of engagement associated with each case study. By viewing these case studies through the micro, meso, and macro perspectives, this ensures that each case study can bring a unique perspective into researching issues of information disorder. The micro/meso/macro perspective is commonly used in sociology and economics (Jepperson and Meyer, 2011) but it has been previously used in social media research, for example as shown in Figure 4-1. Bruns and Moe (2014) associate the 'macro' with 'ad hoc publics', the meso with 'personal publics' and the micro with the 'interpersonal'. In this thesis, The Momo Challenge case study predominately reflects the micro perspective as it focuses on moral panics spread among private and personal communication such as WhatsApp chats (the personal

[3] More information about First Draft's aims can be discovered here: www.firstdraftnews.org/about/

[4] https://www.youtube.com/about/copyright/fair-use/#yt-copyright-resources

[5] https://developers.google.com/youtube/terms/api-services-terms-of-service

perspective) that then reached the macro perspective due to going viral on public platforms such as YouTube (the global perspective). The Notre Dame Fire case study predominately represents the macro perspective as it represents global phenomena, in terms of the number of people publicly engaging in the event on public platforms like YouTube and Twitter and attempting to spread moral panics globally (Bolíbar, 2016; Jepperson and Meyer, 2011), but similar to the Momo Challenge, communication surrounding this case study can also be viewed through the meso and micro perspectives, depending on the social media platforms used to communicate information about these events. In all the micro/meso/macro perspective aids in the explanation of the virality of these case studies and how this virality is platform specific.

The data from these case studies was collected across differing date ranges due to when the data collection occurred (in 2019) and the amounts of data available. The Momo Challenge reflects the historical unfolding of an event, in the sense that it gathers data from over a one-year period. In contrast, the Notre Dame Fire focuses on how breaking news quickly evolves in a short period of time, by only focusing on videos released in a 24-hour period. These case studies illuminate how information disorder spreads in varying time constraints.

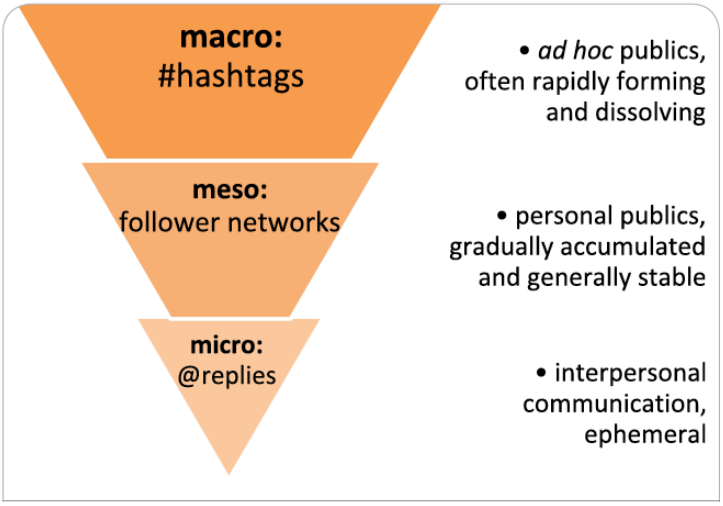


Figure 4-1- Micro, Meso and Macro Levels in Twitter Research (Bruns & Moe, 2013, p.20)

4.2.1. Organisation of Data

Due to the large volume of multimodal data considered in this thesis, data management was critical. The data is divided into four categories: comments, transcripts, videos, and screenshots. For each sampled video, these four forms of data were collected and stored in separate folders. In terms of organisation, each video was given a specific ID and screenshots were saved according to their frame sequence. All videos and their respective metadata were also downloaded. The specifics regarding data collection will now be explained.

4.2.2. Collecting Videos with YouTube Data Tools

YouTube Data Tools (Rieder, 2015) was used to collect YouTube video data for each of the three case studies, that will be detailed later in this section. YouTube Data Tools is a software program designed for researchers, that uses YouTube's Application Program Interface (API) to extract information from YouTube databases.[6] The tool was used to track videos related to the key topics analysed, as well as to determine the videos have with the highest view counts that were most relevant to my search queries. The tool has been used previously to explore YouTube's "ranking cultures", suggesting that YouTube content that appears in the top 20 search results is not just only influenced by 'popularity metrics' but also platform features and sociocultural issues (Rieder et al., 2018).

Identifying popular YouTube videos was important to this study because my research questions seek to understand instances of information disorder that have a broad impact. Whilst previous research has focused on instances of information disorder that have only affected a small number of people, or have had limited spread in comparison to truthful accounts (Lazer et al., 2018; Zubiaga et al., 2015), my study focuses on instances of information disorder on social media that have been viewed by a large

[6] Detailed in the YouTube Data Tools FAQ section: https://tools.digitalmethods.net/netvizz/youtube/faq.php

percentage of people, compared to truthful accounts of the chosen news story. YouTube itself can be likened to a search engine, since many people use it to search for information (Hanson and Haridakis, 2008) and news videos (Peer and Ksiazek, 2011). The moral panic dimension of my research aims also requires that the data be widespread enough to be defined as a moral panic (Cohen, 1972).

The 'Video List Module' *(as shown in Figure 4-2)* was used to collect YouTube video data based on my selected search queries and specified date ranges. The YouTube video data was ranked according to view count. Determining the search criteria for my study, involved considering when each case study was searched for the most and the top key words that were used. This was done using Google Trends[7], which allows a user to see the popularity of terms that were searched for on Google or YouTube within specified date ranges, as well as showing related queries.

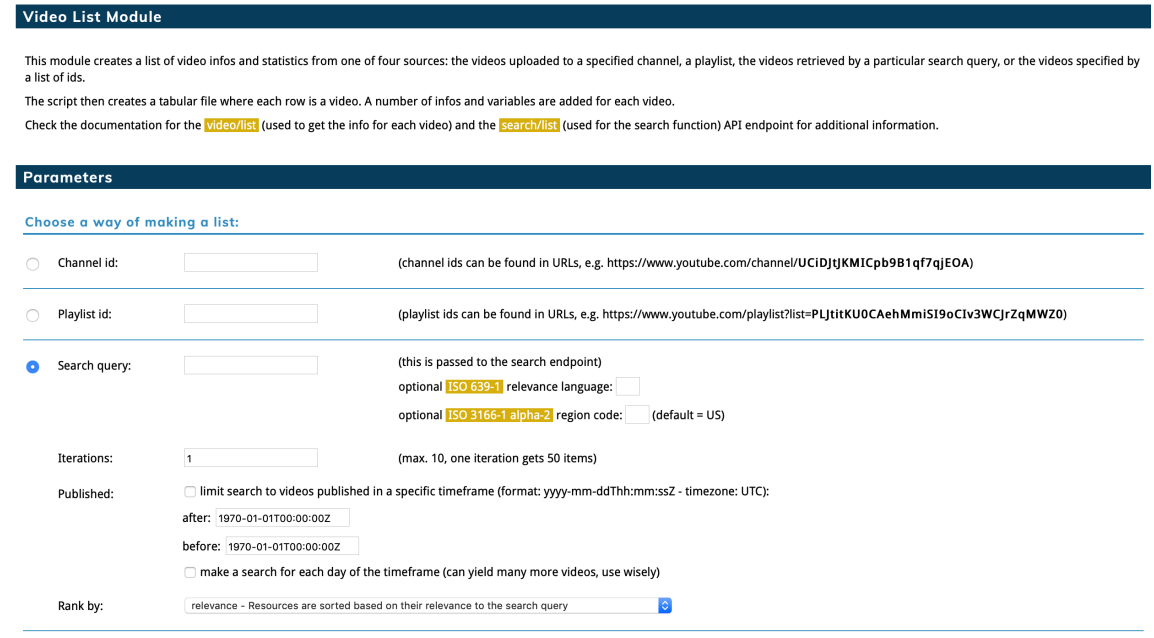


Figure 4-2 - YouTube Data Tools Video List Module

[7] Google Trends: https://trends.google.com/trends/?geo=US

The output of the video list module is in .tab delimited form, that was converted into an excel document. The following information was provided:

- Position (in Youtube List)
- Channel ID
- Channel title
- Video ID
- Date of publication
- Date published at SQL (Structured Query Language)
- Video Title
- Video Description
- Video Category
- Duration
- Duration in seconds
- Dimension
- Definition
- Caption
- Thumbnail link
- Licensed Content
- View Count
- Like Count
- Dislike Count
- Favourite Count
- Comment Count

The most relevant categories for my research were the Video ID, Date of publication, Video Title, Video Description, Video Category, View Count and Comment Count. YouTube Data Tools did not perfectly show the most relevant results related to my search query, so I still needed to sort through the data and eliminate videos that were not in English, as this is beyond the scope of my study, as well as videos that are not related to the topics of my case studies. The next section will provide a description and outline the specific data collection strategies for each case study.

4.2.3. Case Study 1: The Momo Challenge

The Momo Challenge refers to a moral panic shared by various news outlets and concerned parents. It falsely claimed that children were receiving threatening messages via WhatsApp with an image of a frightening figure. Some media outlets also reported that footage of the frightening figure was being spliced into children's YouTube videos such as "Peppa Pig" videos[8]. The Momo challenge has now been discredited[9] and YouTube has stated that it did not discover any spliced videos[10]. Therefore, the Momo challenge is now referred to as an "internet hoax"[11].

In order to further explain how the Momo Challenge developed and its cultural impact, the Wikipedia page for the Momo Challenge provides some interesting insights. The title of the Wikipedia page is now the "Momo Challenge hoax" explicitly indicating the falseness of the Momo Challenge. However, when the first Wikipedia page on the Momo Challenge was created on 21st August 2018, the title of the page was simply "Momo Challenge" and was described as a "form of cyberbullying" rather than an internet hoax. This original Wikipedia article, states that the Momo Challenge:

> ...gained the public's attention in July, 2018, when it was noticed by a youtuber with a large following.

This article also spreads false information about the Momo challenge in its introduction:

---

[8] See: https://www.thesun.co.uk/news/8515907/momo-challenge-fortnite-youtube-instagram-self-harm/

[9] https://www.theguardian.com/technology/2019/feb/28/viral-momo-challenge-is-a-malicious-hoax-say-charities ; https://www.theatlantic.com/technology/archive/2019/02/momo-challenge-hoax/583825/ ; https://www.snopes.com/news/2019/02/26/momo-challenge-suicide-game/

[10] YouTube statement: https://support.google.com/youtube/thread/1917881?hl=en

[11] Discussion of experts falling for the momo challenge internet hoax: https://theconversation.com/momo-challenge-shows-how-even-experts-are-falling-for-digital-hoaxes-112782

> Its spread and the suicide of a 12-year old in Argentina are stoking fears that this could be the next Blue Whale.

The details in this early Wikipedia article match the sentiments shared by many of the YouTube videos that will be analysed in this chapter. The development of the Wikipedia page provides some contextual insight into how the Momo Challenge has been perceived over time. However, whilst many of the YouTube videos analysed in this thesis aligned with the views stated in the earlier version of the Wikipedia article, the article has evolved over time, to align itself more with news coverage provided on the Momo Challenge by outlets such as the Guardian, Vox, New York Times, The Atlantic and BBC that directly call out the Momo Challenge as a hoax.

In the talk page (the name Wikipedia gives for the discussion forum that is linked to a specific Wikipedia page and is a space where writers and editors discuss the writing/editing process for a Wikipedia page) for the "Momo Challenge Hoax" there is discussion about information versus speculation. Some Wikipedia users try to assert that there are suicides linked to the Momo Challenge but are rebutted by other users stating that there is no evidence:

> We should stick with confirmed facts (police investigating vs confirmed link to Momo challenge).

> Grieving parents, whilst under considerable strain and possibly talking to journalists shortly after the event are not reliable sources.

> We've had that debate with the Blue Whale Challenge page last year, with large numbers of complaints from citizens that ended up being unfounded - real, heartbreaking suicide cases but nothing to do with Blue Whale.

This talk page, where discussion of 'evidence' takes precedence, shows the evolution of the Momo Challenge Wikipedia article. In response to many of these discussions in the Wikipedia talk page, the introduction of the 'Momo Challenge hoax' Wikipedia page was changed significantly, explicitly stating that the Momo Challenge is a hoax:

> The "Momo Challenge" was a hoax and an Internet urban legend about a nonexistent social media challenge that was spread on Facebook and other media outlets.

For the Momo Challenge case study, based on Google Trends data (Figure 4-3), YouTube videos on the subject matter started appearing on 15th July and the last video uploaded based on the Momo Challenge was on 11th August 2019. The most common search terms used were: *momo challenge, momo peppa pig challenge, peppa pig, momo peppa pig, peppa pig momo, momo challenge videos, momo challenge horror videos*. Based on this data, these search terms and date range were selected for the Video List Module on YouTube Data Tools. The videos were ranked according to "viewCount", sorted from highest to lowest number of views.

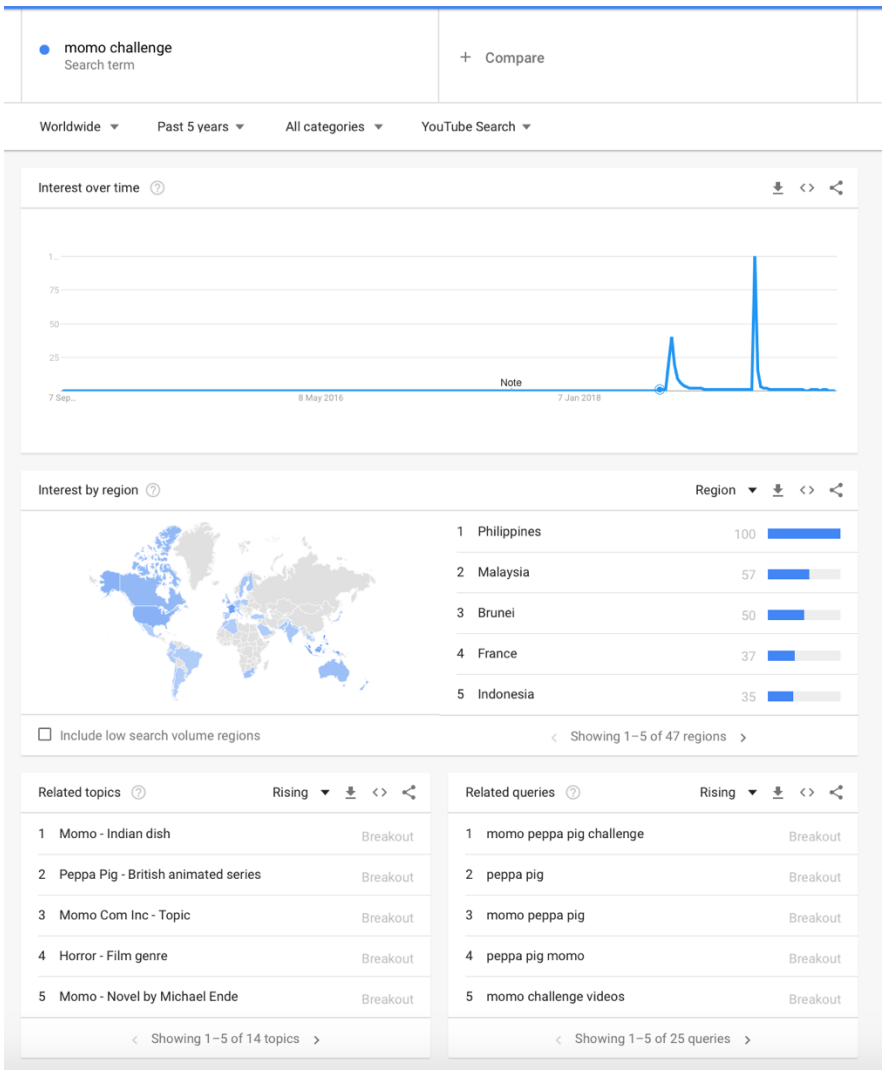


Figure 4-3 - Momo Challenge Google Trends Data

Videos in languages other than English were included in the list. Therefore, I needed to manually sort through the videos and determine which videos were predominantly in English (i.e. actual video content was in English, not just the title, description or captions), as analysing videos in languages other than English is beyond the scope of this study. Due to the number of videos that were also gathered by the YouTube API, I limited my analysis to videos with more than 10 000 views. After this process I was left with 195 videos. These 195 videos were sorted into macro-genres, based on their perceived purpose e.g. news reporting, commentary, entertainment or advertising, and level of deceptiveness; satire, unintentional misinformation, intentional disinformation or unclear.

Once the 195 videos were selected, a grounded theory approach (Strauss and Corbin, 1997) was used to sample the videos that would be analysed in detail. This meant that relevant discursive patterns were searched for in the data, until I reached a point of saturation, where no new patterns could be detected. A range of videos needed to be selected that represented each respective genre identified by the grounded theory approach: news reporting, entertainment, commentary, educational, and click-bait/attention seeking. As a detailed analysis would be conducted on the videos, a qualitative approach was favoured. Videos needed to be selected that would represent a selected genre, received more than 10,000 views, and had an active comments section. Videos that had 'comments disabled' were not chosen for closer analysis.

Macro-Genre + Level of Deceptiveness	**Channel Number (anonymised)**	**Video Title**
News reporting (misinformation)	1	Disturbing 'Momo Challenge' suicide game concerning schools, parents
News reporting (misinformation)	2	Parents warn about potentially deadly 'Momo Challenge' online
News reporting (misinformation)	3	Who's Behind the 'Momo Challenge?'
Entertainment (conspiracy)	4	Real Life Momo Challenge (Very Scary) Calling Momo Caught on Camera
Entertainment (misinformation)	5	"THE MOMO CHALLENGE" Peppa Pig ORIGINAL VIDEO REACTION (PARENTS PLEASE WATCH) *Momo is dead*
Entertainment (misinformation)	6	YOUTUBERS REACT TO MOMO (Scary Meme or Hoax?)
Commentary (Summarising the momo challenge – mix of facts and misinformation)	7	Exploring the Momo Situation
Commentary (misinformation)	8	Another Boy Died Because of the Momo Challenge
Commentary (misinformation)	9	Momo Challenge Peppa Pig Original Video: 'Suicide game' on YouTube, KIDS, Fortnite

Educational (Addressed to Parents)	10	*TAKE NOTICE* The Momo Challenge & Peppa Pig Videos Are NOT FAKE!
Educational (Explaining how the momo challenge works)	11	How the Momo Challenge uses Psychology to Kill
Educational (Addressed to Parents)	12	The Momo Challenge Explained
Clickbait/attention-seeking	13	MOMO CHALLENGE EXPOSED | DON'T DO IT | DO NOT CALL PERIOD | WHATASAPP GAME HOAX WARNING
Clickbait/attention-seeking	14	EXPOSING THE MOMO CHALLENGE *REAL VIDEO EVIDENCES*
Clickbait/attention-seeking	15	ITS REAL!!!! MOMO CHALLENGE REACTION!! (IM SCARED)

Table 4-1 - A Summary of the 15 videos selected for the Momo Challenge Case Study

Overall, by adopting an approach that used macro-genres as an initial selection criterion for inclusion in the dataset, a wide selection of Momo challenge videos could be selected. In terms of the research question, by identifying that there are a range of different macro-genres in the video dataset, the study was able to consider how instances of problematic content can be different depending on the genre of a video.

4.2.4. Case Study 2: The Notre Dame Fire

The second case study dealt with the Notre Dame Fire, spanning the timeframe of the fire on 15th April 2019 and the immediately unfolding news coverage. Due to extensive media attention and the initial uncertainty about the cause of the fire, the Notre Dame Fire was subject to significant amounts of misinformation by those unknowingly spreading false information, disinformation by those seeking to blame their political opponents for the fire, and conspiracy theories by those ranging from those with a political agenda to those with purely a fascination in challenging mainstream media.

Coverage of the event started appearing on 15th April. The 16th April was selected as the end date range, in order to focus on the breaking news coverage of the event, rather than changing perceptions of the fire overtime. The selected key words for the search query were: *Notre dame, notre dame paris, notre dame cathedral, notre dame de paris, notre dame fire,* based on the most common google web searches (see Figure 4-4). Some videos were removed from the dataset because they were not in English. The analysis was limited to videos with more than 10,000 views, resulting in a set of 272 videos which were then categorised based on whether they represented factual news reporting, conspiracy theories, or disinformation.

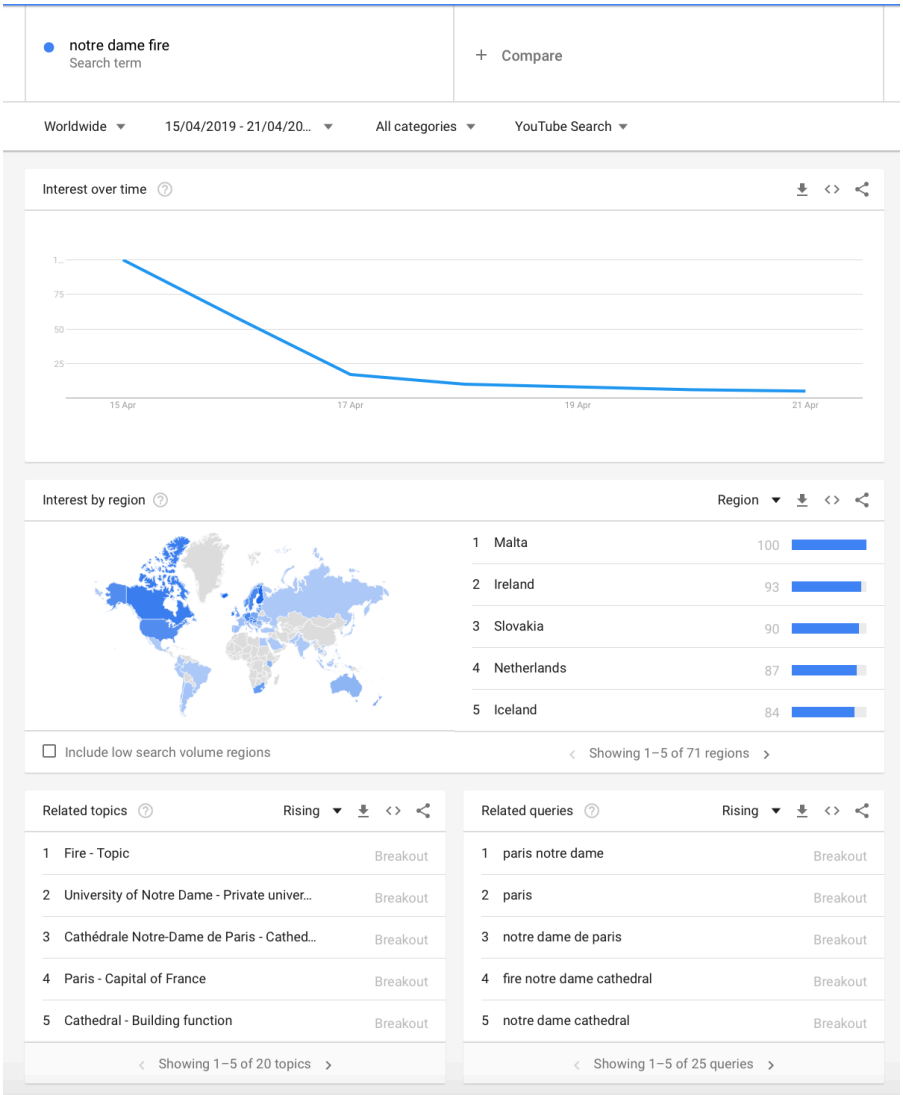


Figure 4-4 - Notre Dame Fire Google Trends Data

The grounded theory sampling approach applied to the dataset revealed two main genres: commentary and news report. Levels of deceptiveness varied among disinformation, hate speech and conspiracy. A summary of the videos selected for closer detail is provided in Table 4-2.

Macro-Genre + Level of Deceptiveness	**Channel Number (anonymised)**	**Video Title**
Commentary (disinformation and hate speech)	1	The Notre Dame Fire
Commentary (disinformation and hate speech)	2	What They're NOT Telling You About The Notre Dame Fire
Commentary (disinformation and hate speech)	3	"We wish more fire upon you" – Muslim world reacts to Notre Dame tragedy
News report (disinformation)	4	Notre Dame Cathedral fire: Buzzfeed FAILS with "anti-Muslim narrative" | Martina Markota
Commentary (conspiracy)	5	Mystery figure at Notre Dame cathedral fire
News report (misinformation)	6	Notre Dame Cathedral Fire: Suspicion After Hundreds of French Churches Vandalised | Martina Markota
Commentary (conspiracy)	7	Fire At Notre Dame Follows Wave Of Church ATTACKS – Will They Blame The Yellow Vest Movement?
Commentary (conspiracy)	8	Notre Dame Fire: Globalist False Flag to Trigger WWII?
Commentary (conspiracy)	9	Notre Dame Paris Fire INSIDE JOB, Destroying Tartarian Art by Burning
Commentary (conspiracy)	10	The Notre Dame Cathedral Fire | A Planned, Deliberate Event
Commentary (UFO conspiracy)	11	Now 3 UFOs Filmed At Notre Dame Fire! OMG!
Commentary (conspiracy)	12	NOTRE DAME FIRE: Suspicious Activity on Roof

Commentary (conspiracy)	13	THE SCARY TRUTH ABIOUT THE NOTRE DAME FIRE THAT NO ONE IS TALKING ABOUT…
Commentary (UFO conspiracy)	14	Breaking: "Nostradamus Predicted Paris Notre Dame Would Burn" / WW II
Commentary (conspiracy)	15	The Notre Dame Cathedral Fire was Arson
Commentary (conspiracy)	16	What's the truth abou the Notre Dame Cathedral Fire?

Table 4-2 - A Summary of the 15 videos selected for the Notre Dame Case Study

Overall, the majority of videos selected represent a range of different conspiracy theories regarding what caused the Notre Dame Fire and challenging the factual information indicating that the Notre Dame fire was not deliberate.

4.2.5. Collecting Comments with YouTube Comment Suite

The comments for each video were collected with the software program YouTube Comment Suite. YouTube Comment Suite uses the YouTube API (Application Program Interface) to extract the comments from any selected video. As the YouTube videos analysed still have active links online, there is no time limit restricting users commenting on selected videos. Nonetheless, in terms of the data collection process in this thesis, an end date had to be selected for each video, in order to have a stable amount of comments to analyse. The date range for the extracted comments was decided upon by following the Google trends data as outlined in the previous sections for each case study. For case study 1 the end date for all the comments was the 11st August 2019, the date when videos stopped getting produced on YouTube about the Momo challenge, and the challenge stopped trending. For case study 2, the end date was 1st January 2020, as many comments were posted over a much longer time frame, reflecting continual interaction with the news story. YouTube Comment Suite also provided graphs on each comment thread, so one can see when most comments were

posted. An example is provided in Figure 4-5. With most videos, there were peaks in comments around when the video was first released, and comments usually dropped off after 30 days. This matches research by data analytic firms that state the half-life (time it takes for a video to reach 50% of its lifetime audience) of YouTube is 6+ days whilst after 20+ days a YouTube video reaches 75% of its lifetime audience (Liebmann, 2018; Yarow and Angelova, 2010).

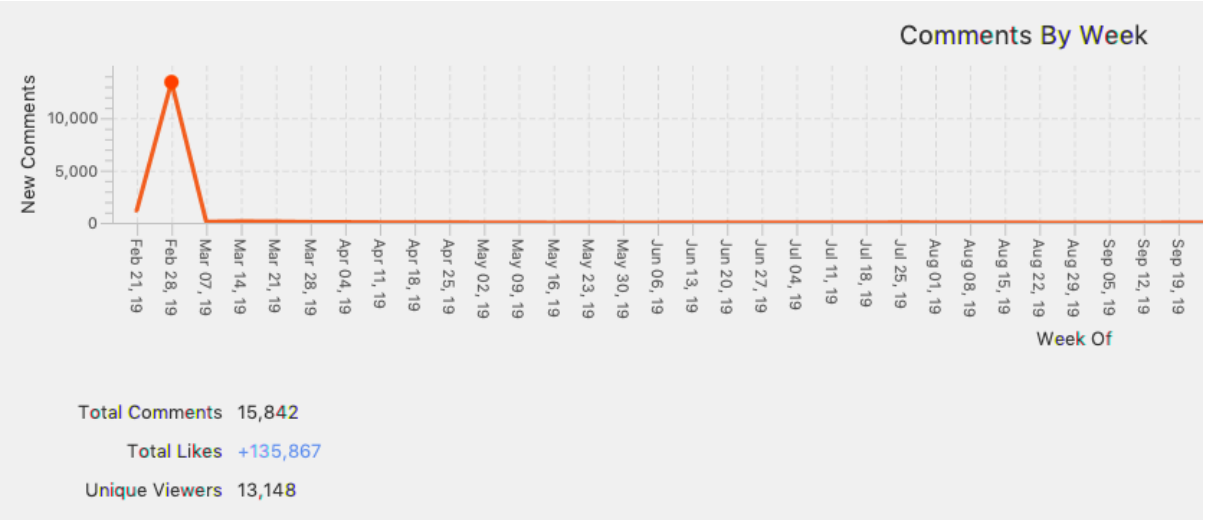


Figure 4-5 - YouTube Comment Suite Screenshot

4.2.6. Corpus Linguistics and Grounded Theory as Data Collection Methods

Once all the comments were collected, a mixed corpus linguistics and grounded theory approach was used in order to collate a sample of comments that could be analysed in relation to the research question. Using the concordance software *AntConc* (Anthony, 2014), a list of the most common words across the entire comment corpus for each case study was compiled, in order to determine if there were any trends across all the comment threads. By selecting the most common words across the entire dataset, this allowed commonalities to be discovered among the videos and to effectively compare how users bonded over certain topics. Common words were selected over n-grams because the sample of n-grams did not generate many comments. In comparison, common words generated a larger list of comments. The results from this corpus

linguistics approach formed the basis for a closer discourse analysis, utilising the appraisal framework to analyse the role of evaluative meaning in language (Martin and White, 2005) and affiliation analysis (Zappavigna, 2018) to explore how people bond around certain values.

Completing a detailed analysis of YouTube comments can be a difficult task, due to the number of comments readily available. Thus, for this project it was essential to create a sampling strategy so that comments could be analysed in more detail. Previous research regarding YouTube comment sampling, has adopted strategies such as random sampling with an inclusion criterion (Jeon et al., 2018), first 1000 comments of a video (Thelwall et al., 2012) and profile-sampling (Ernst et al., 2017). For this thesis, a theoretical sampling process was selected. Theoretical sampling is a part of the grounded theory method and is defined as:

> '...the process of data collection for generating theory whereby the analyst jointly collects, codes, and analyzes his data and decides what data to collect next and where to find it, in order to develop his theory as it emerges' (Glaser Barney and Strauss Anselm, 1967: 45; Hadley, 2017: 33).

The grounded theory element to this sampling strategy involved selecting comments that featured in the corpus linguistics sampling strategy. As the purpose of this part of the thesis and research question is to understand how people affiliate around information disorders, in this case it is more crucial to find instances of common themes across the dataset, rather than the most popular comment.

For The Momo Challenge case study, based on the commonalities among themes in this dataset, the most common word was selected, which was 'Momo'. Concordance lines across all the individual comment threads were examined. An example is shown in Figure 4-6. Some videos still had more than 100 comments featuring the word 'Momo', so in this case, theoretical sampling was used to select a more manageable range of comments. This involved dividing comments initially into three categories:

1. whether users expressed that the Momo challenge was real,
2. whether users expressed the Momo challenge was false,
3. comments that did not comment on the veracity of the Momo challenge.

As the research question is focused on understanding how people bond around information disorders, the comments that were debating whether the Momo challenge was true or false were selected. More particular themes within these comments were then identified, which included serious explanations of why Momo was true, recounts of hearsay in regard to people affected by the Momo challenge and comedic interpretations of Momo as real, so that a smaller dataset of roughly 50 comments for each video across a range of themes could be created for the appraisal and affiliation analysis.

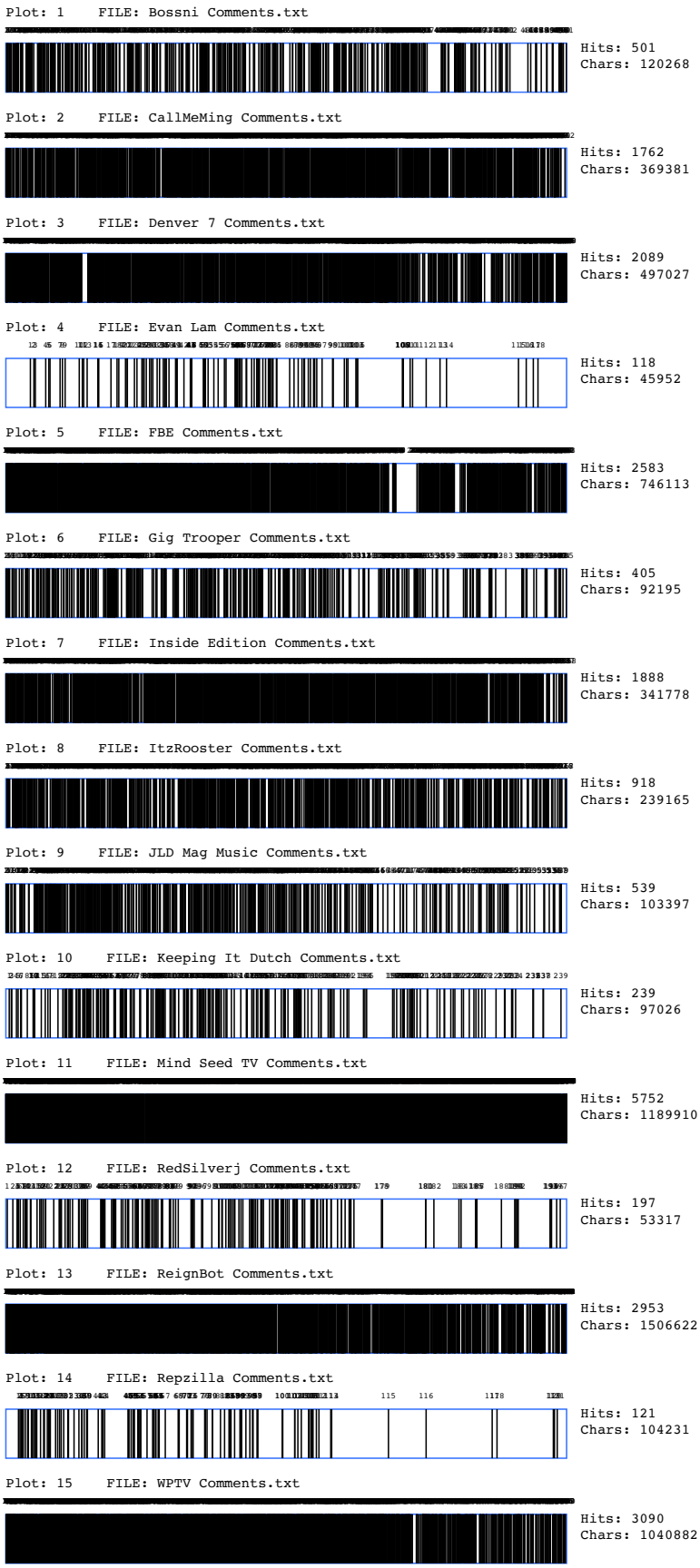


Figure 4-6 - Frequency Lists of 'Momo' Appearing in a Corpus

For The Notre Dame Fire case study, all the videos consisted of more than 100 comments. Again, the most common word was selected (which was "fire"), from which the 50 comments were sampled for each video. However, there was one video in this dataset that did not have a keyword that featured more than 100 times. In this case, the most liked 100 comments were selected instead.

4.2.7. Collecting YouTube Screenshots and Transcripts

For each video, transcripts and visual data needed to be collected. Each video was downloaded from YouTube as an MP4 file to maintain a permanent record. YouTube's automatic transcription tool provided a simple, initial way to store transcription data. However, YouTube's transcription service is not perfect, therefore the automatic transcriptions were edited in order to provide greater accuracy. This included editing the transcripts to reflect a change of speakers, correct spelling, and punctuation mistakes, and to alter the layout of the transcripts so it was easier to read.

Screenshots of each frame were essential for a multimodal analysis of the videos. Therefore, after converting the videos into MP4, the open-source software VLC player was used to capture automatic screenshots. This software recorded a screenshot of each 20 second frame of a video. Each screenshot was labelled according to its timestamp. Thus, a large batch of screenshots was generated, with over 1000 screenshots per video. In order to analyse a manageable set of screenshots, the screenshots were sampled in terms of a grounded theory approach. This meant that every time a frame changed to a new scene or genre stage (e.g., from a YouTuber talking to the camera to a full screen image of a tweet; or from an explanation of an event to an interview) it was sampled. The sampled screenshots were included in an excel spreadsheet tab for each video, and each screenshot was labelled according to its genre stage and analysed according to the multimodal analysis detailed in Section 4.3.6. On average, each video according to this method had at least 50 screenshots (if it was only a two-minute video) or 200 screenshots if the video was five minutes or longer (refer to *Volume II: Appendices* to get a sense of all the data). Ultimately, the number of screenshots for analysis depended on the amount of dynamically changing scenes in the video e.g., a video that just focused

on a YouTuber talking to the camera had fewer screenshots than a video that frequently switched from the YouTuber to another sampled video or full screen screenshot or tweet. By sampling images according to the aforementioned strategy, this ensures that the research question is directly answered, in terms of being able to conduct a close analysis of screenshots with multiple layers of meaning.

4.3. Analysis of Data

The data analysis involved a multifaceted approach, with different forms of analysis conducted on the various kinds of YouTube Data collected in the study. These different forms of YouTube data are summarised in Table 4-3.

Data Type	**Explanation**	**Instances**
YouTube video metadata	Titles, channel names, descriptions of YouTube videos, and a list of closely related videos obtained via YouTube Data Tools	Data obtained for 30 videos: video descriptions 6120 words in total
Transcripts of YouTube videos	With timestamps derived from YouTube's automatic transcription service	Data obtained for 30 videos: 32 732 words in total
Comments on YouTube videos	Obtained via YouTube Comment Scraper tool	Data obtained for 30 videos: 1500 initiating comments in total and 1461 replies in total
Visual content of YouTube videos	Obtained by taking screenshots of videos	Data obtained for 30 videos: 1674 screenshots in total

Table 4-3 – YouTube Forms of Data Summary

A summary of the data analysis conducted on each form of data and why it was conducted is briefly provided here:

1. **Content analysis of video metadata** – This was a 'soft-eye' approach categorising videos into broad text types based on titles and descriptions provided. This approach provided the basis for understanding the different sorts of deceptive communication that exist and how these videos attract a user's attention.

2. **Appraisal, Affiliation, and Legitimation analysis of transcripts** – A combined appraisal, affiliation and legitimation analysis was conducted on the transcripts of video data. This revealed key evaluative meanings in discourse, the social bonds that YouTubers promoted in their videos, and the legitimation strategies YouTubers used in order to legitimate these social bonds. In addition, patterns in the results discovered from this analysis revealed new ways of understanding how genres of information disorder are constructed.

3. **Appraisal, Affiliation, and Legitimation analysis of YouTube comments** – A corpus linguistics approach was used to sample comments based around the most common word featured in the corpus. An appraisal analysis was then conducted on the developed comment corpus. From this, the coupling of ideational and attitudinal meaning was analysed, in order to understand what concepts people are affiliating around and the particular affiliation strategies use was identified. The role of this affiliation was then considered in relation to the legitimation framework, in order to analyse how the key social bonds that users shared were legitimated.

4. **Multimodal analysis of visual video content** – Screenshots of the videos were taken and timestamped (based on YouTube's automatic transcription function). In particular, attention was focused on the images shown at the same time as the deceptive or problematic textual message unfolds, and how images contribute to the deceptive and problematic impact of a YouTube video. Currently, not much current research exists in relation to social semiotic multimodal interpretations of YouTube videos. Current theorists in the area of social semiotic interpretations of moving images include Feng and O'Halloran (2013) who have

applied a mixed SFL and cognitive linguistics approach to analysing moving images, and Bateman (2008) who has the considered the multimodal document in relation to static images, and has later extended this to moving images as well (Bateman, 2013). In terms of multimodal social semiotic approaches, Zappavigna (2019) has examined YouTube videos by focusing on language, gesture and visual frame. Benson (2016) has applied a multimodal social semiotic approach to the analysis of YouTube video page layouts and interaction in comments but has not analysed the videos themselves. My approach to analysing the visual content of YouTube videos involved developing a framework that takes a metafunctional approach to multimodality, a similar method to work by Bednarek (2014) and Kress and Van Leeuwen (2006) as will be further explained in Section 4.3.6.

4.3.1. Content Analysis of Video Metadata

A more detailed explanation of the data analysis process will now be provided. A content analysis of the video metadata extracted with YouTube Data Tools (as detailed in the previous section) provided the basis for determining which videos to focus my analysis on. After I sorted through the collection of excel spreadsheet data, I created a Microsoft word document with the title and description of each video. I then imported this document into the UAM Corpus Tool[12], a software program that is designed for the annotation of textual corpora. Based on the titles and description of each video, the videos were initially arranged into macro-genres on the basis of what I perceived as the purpose of each video in terms of its genre stages e.g., entertainment, advertising, opinion or news reporting. I also analysed the level of deceptiveness of each video based on the video descriptions e.g. is most information factual, satire, unintentional misinformation, intentional disinformation, or unclear. My knowledge of the deceptiveness of each video was derived from contextual knowledge regarding the news coverage and history behind the case studies.

[12] http://www.corpustool.com

This content analysis of video metadata helped me determine which videos to sample for each case study for close analysis using the methods explained in the sections which follow on the appraisal, affiliation and legitimation analysis. In order to select the videos that I would analyse in detail, my research adopted a grounded theory approach (Strauss and Corbin, 1997). This means that, with my research aim in mind (understanding the variety of information disorders online), I selected a video that represented each type of macro-genre I identified. Sampling ceased once I reached the 'saturation of description' point, meaning that the collection of additional videos did not modify the number of different macro-genres and information types I had identified (Strauss and Corbin, 1997).

4.3.2. Appraisal Analysis on Transcripts and Comments

An appraisal analysis was conducted on the transcripts and comments of the selected videos. This is outlined in Table 4-4. As this table highlights, the 15 videos for each case study featured numerous data that was manually analysed.

Case Study	**Transcript Data**	**Comment Data**
The Momo Challenge (15 videos)	14 426 words	750 initiating comments 507 replies
The Notre Dame Fire (15 videos)	18 306 words	750 initiating comments 954 replies

Table 4-4 – Summary of Text Data

The Appraisal framework (Martin and White, 2005) consists of three key dimensions: attitude (feelings and judgements of behaviour), graduation (judging how strong a feeling is), and engagement (the voices around opinions in discourse). There are sub-types within these three key dimensions:

- **ATTITUDE** (expressing emotion or opinion) as AFFECT (feelings) JUDGEMENT (ethically assessing a person or behaviour) or APPRECIATION (valuing an object or phenomenon).
- **GRADUATION** (up or downscaling attitude) as FORCE (via intensity) or FOCUS (via prototypicality).
- **ENGAGEMENT** (managing the play of textual voices) as MONOGLOSS (expressing only one voice) or HETEROGLOSS (expressing multiple voices).

The appraisal system network is shown in Figure 4-7.

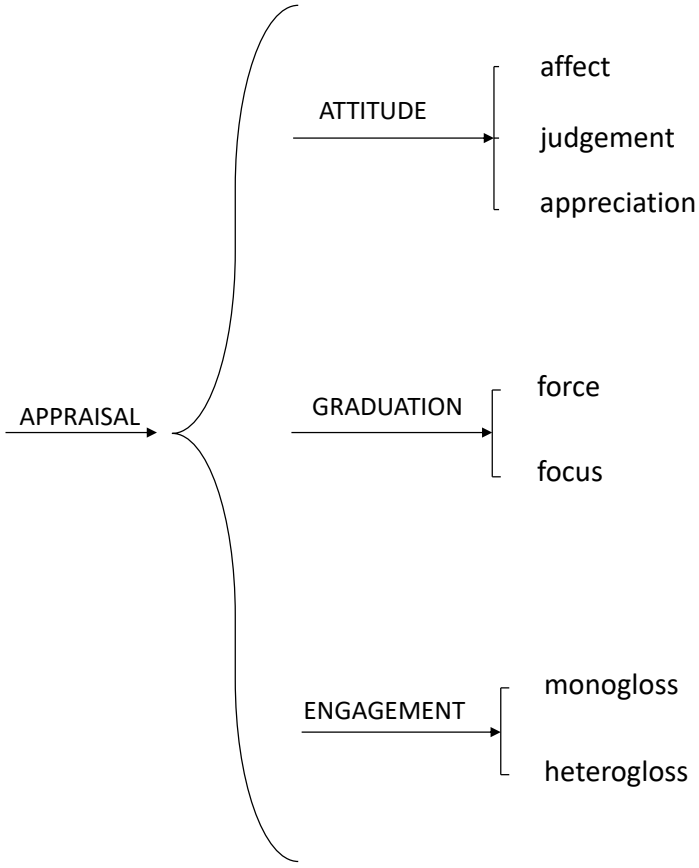


Figure 4-7 - Appraisal Framework (Martin & White, 2005)

There were also further sub-systems within AFFECT, JUDGEMENT and APPRECIATION that were explored. Whilst AFFECT and JUDGEMENT are reserved for the feelings and

judgements directed towards humans, APPRECIATION refers to evaluations directed towards objects and things. These further sub-systems are shown in Table 4-5.

Appraisal			**Explanation of Sub-Type**
Affect	Un/Happiness		Moods of feeling happy or sad
	In/Security		Feelings of peace and anxiety
	Dis/Satisfaction		Feelings of achievement and frustration
	Dis/Inclination		Feelings of fear and desire (irrealis affect)
Judgement	Social Esteem	Normality	How special?
		Capacity	How capable?
		Tenacity	How dependable?
	Social Sanction	Veracity	How honest?
		Propriety	How far beyond reproach?
Appreciation	Reaction		Impact and quality
	Composition		Balance and complexity
	Valuation		Was it worthwhile?

Table 4-5 - Attitude System Sub-Types (adapted from Martin and White, 2005)

In this thesis I was interested in understanding values as linguistic couplings of ideational and attitudinal meaning. This requires considering the targets of appraisal since we "don't after all simply affiliate with feelings; we affiliate with feelings about people, places and things, and the activities they participate in, however abstract or concrete" (Martin, 2008a: 58). The following annotation convention derived from Zappavigna and Martin (2018) was used to mark up the data for 'couplings' of ideation and attitude:

[ideation: <<>>/ attitude: <<>>]

AFFILIATION STRATEGY

The square brackets and the / symbol are used to suggest the fusion of ideation (what is being evaluated) and attitude (the evaluation) to form a value that can be negotiated through the process of affiliation, that will be explained in further detail in the following

section. Examples throughout this thesis also show the ATTITUDE in **bold** and ideation underlined.

Each transcript was coded, to analyse inscribed vs. invoked attitude, and the various systems and sub-systems of attitude, engagement, and graduation. A summary of what has been coded for is provided in Table 4-6.

Type of Attitude	**Attitudinal System**	**Graduation System**	**Engagement System**
Inscribed Attitude	Appreciation - Reaction - Composition - Valuation	Force	Monogloss
Invoked Attitude	Judgement - Normality - Capacity - Tenacity	Focus	Heterogloss
	Affect - Un/Happiness - In/Security - Dis/Satisfaction - Dis/Inclination		

Table 4-6 – Summary of Analytical Coding Scheme

The coding of these transcripts identified whether the majority of videos invoke positive or negative attitudes, if these attitudes are expressed with emphasis, and if monogloss (only one view is expressed) or heterogloss (multiple views are expressed) in terms of the voices in the text. This coding is useful in answering the research questions regarding how to analyse the interpersonal function of information disorders.

4.3.3. Genre Analysis of Transcripts

After the appraisal analysis on the transcripts, I determined if there were any linguistic patterns that suggested different genre stages present in the selected YouTube videos and interpreted this in relation to the findings from the appraisal analysis. A summary of the transcript data analysed is provided in Table 4-7.

Case Study	**Transcript words**	**Number of ideation/attitude couplings annotated**
The Momo Challenge	14 426 words	513
The Notre Dame Fire	18 306 words	784

Table 4-7 – Summary of Transcript Data

An example of how different genres stages were considered is shown in Figure 4-8. Functional genre labels for the videos include:

- **Providing Context** about a Situation
- **Explaining** a situation
- **Describing** a situation
- **Linking** a situation to another situation
- Showing something like a social media post as **evidence**
- **Warning** about a situation

Overall, this analysis reveals the similarities and differences in the construction of the YouTube videos. In addition to the discourse and multimodal analysis, it contributes to answering the research question on what macro-genres of information disorder exist online and how content is created to appear credible.

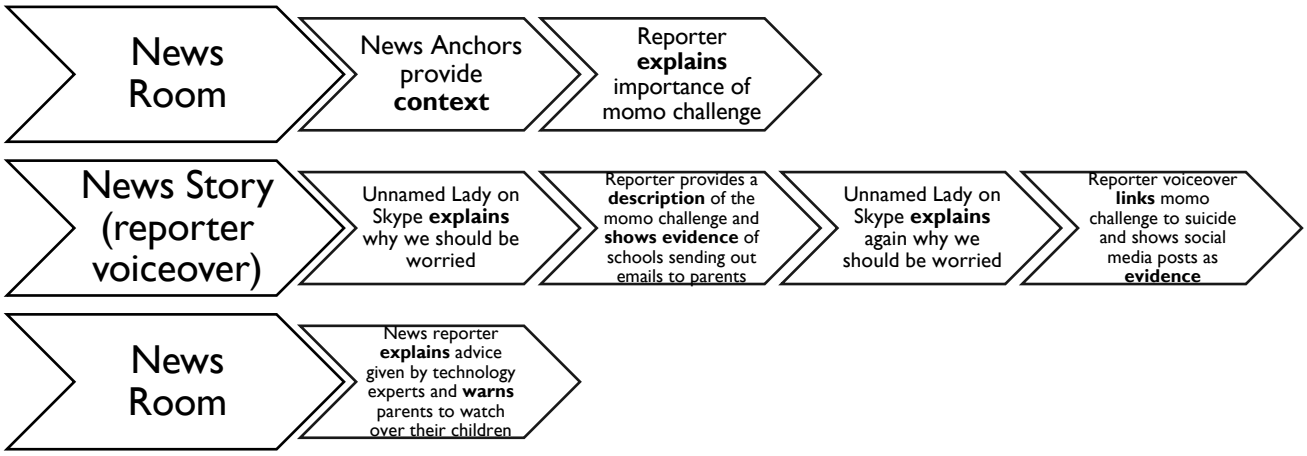


Figure 4-8 – Examples of Genre Stages in a YouTube Video

4.3.4. Affiliation Analysis on Videos and Comments

An affiliation analysis shows in detail how people align or de-align with the social bonds in the videos and comments. This part of the analysis involved applying an appraisal and affiliation analysis to the comment datasets. Whilst appraisal explains the main attitudinal forces at play in terms of the attitudes people are bonding around, an affiliation analysis illuminates how this bonding occurs by identifying the particular affiliation strategies used such as whether a persona is supporting or rejecting a potential bond. The detailed breakdown of the comments analysed is shown in Table 4-8.

Case Study	**Initiating comments analysed**	**Replies analysed**
The Momo Challenge	750 comments in total 50 comments per video	507 replies
The Notre Dame Fire	750 comments in total 50 comments per video	954 replies

Table 4-8 - Summary of Comments Analysis

The entire transcripts of each video (n=30) and sample YouTube comments from the selected videos (n=50) for each case study were analysed. These sampled comments were selected on the basis that they contain the most common reoccurring noun featured in the dataset of each specific case study. YouTube comments with replies were analysed according to the dialogic affiliation network, where users that support a statement for example can 'rally' around a bond, 'adjust' the bond or 'defer' the bond (Zappavigna, 2018). Or if users reject a statement, they can either completely 'dismiss' the bond, or 'oppose' the bond (Zappavigna, 2018). The dialogic affiliation framework is shown in Figure *4-9*.

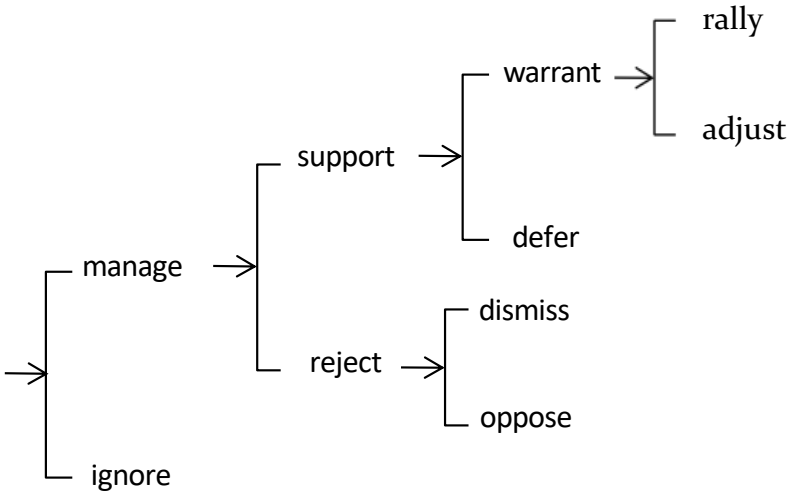


Figure 4-9 – Dialogic Affiliation Network (adapted from Zappavigna, 2018)

For the 'initiating comments' (comments which aren't replies to an initial comment) the affiliation strategy was analysed using a modified version of the communing affiliation system that was originally used by Zappavigna (2018) to explore the kinds of ambient affiliation possible in hashtag use in social media discourse. The revised communing affiliation framework is shown in *Figure 4-10*. This system captures three key affiliation strategies:

- **CONVOKING**: situating a bond through resources that muster together a community, for instance by the use of vocatives such as "*Guys*!" (MARSHAL),

and/or establish its parameters, for instance by naming or referencing a particular community such as "*Here in the Philippines"* (DESIGNATE).

- **TEMPERING**[13]: adjusting a bond through resources that modify the degree of venture, for example through quantification "*100s of*" (FOSTER), or modifying the of scope of venture, for example "*just"* (MODULATE).

- **FINESSING**: assembling a bond through resources that broaden (EMBELLISH) or confine (DISTIL) its intersubjectivity. For example, embellishing would open the bond to various other possibilities offered by a range of voices with differing perspectives ("*I guess"*) whereas distilling would limit this range, often to only one choice ("*it is not*").

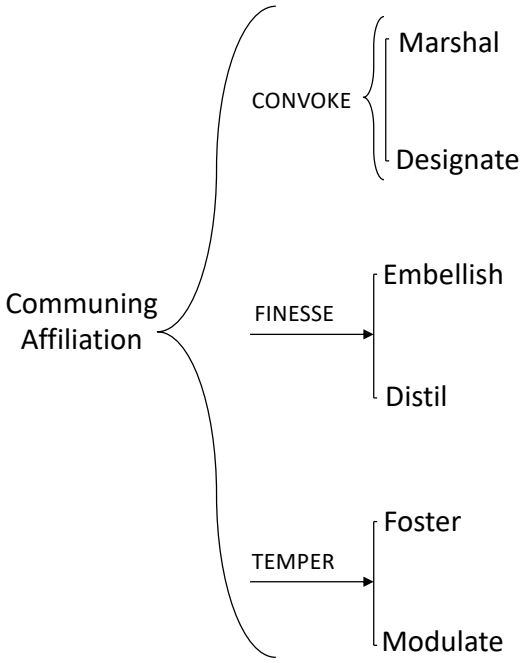


Figure 4-10 – Communing Affiliation Framework (adapted from Zappavigna, 2018)

[13] PROMOTING was renamed TEMPERING to account for both upscaling and downscaling resources of the kind not possible in hashtags (which inherently make a bond more visible).

By understanding the affiliative strategies used in the comment corpus, the most common patterns in the affiliative strategies can be identified, along with how these strategies are used recurrently by personae (as will be explained more in the next section) as they table bonds as various ideation-attitude couplings. Overall, an affiliation analysis represents an empirical approach to understanding how communities of shared values are discursively constructed and will answer the research question regarding how users bond around information disorders and their motivations for doing so in terms of the social bonds they most closely align with.

4.3.5. Identifying Personae in Comments

In this thesis, personae are assigned to different categories based on the bonds that they share, an approach similar to Knight's (2010b) concept of discursively formed communal identity. This section will summarise how textual personae were identified in the comment dataset (building on the explanations of appraisal and affiliation stated earlier in this chapter) and the annotation strategy for initiating comments (the comments that were labelled according to textual personae).

Textual personae in this thesis, drawing on the SFL meaning of personae (refer to *The Literature Review* for research previously conducted into personae from different linguistic perspectives) is based on the key social bonds that are linguistically enacted in discourse. These social bonds (the broader values communities align around) are based on the generalisation of the ideational and attitudinal coupling analysis (annotated as [ideation:.../attitude...]) conducted on the comments. In other words, the social bonds reflect the macro-level attitudes from the ideation-attitude analysis (i.e., what is at stake in this evaluation and what is the broader entity or phenomenon being evaluated). In order to work out these social bonds contextual knowledge is needed, therefore this would be a difficult process to automate.

The process of defining textual personae was based on a grounded theory method, whereby after the appraisal, affiliation, and social bonds analysis was conducted on the

comments, patterns were searched for in the data (social bonds) until all possible patterns and groupings were exhausted. The personae labels were based on the predominant linguistic strategy of the grouping of social bonds e.g., comments that held social bonds that all expressed scepticism of the veracity of a video (FAKE VIDEO BOND) belonged to the Sceptic textual personae.

To provide another example, this comment has been annotated according to the appraisal, affiliation, and social bonds analysis:

I knew it was <u>arson</u> the moment I saw it on fire.
[ideation: arson/ attitude: invoked positive VALUATION]
<u>No MSM report</u> **will change my mind**.
[ideation: MSM/ attitude: invoked negative VERACITY]
Finesse: Distil (I knew) ↘ DELIBERATE FIRE BOND
Finesse: Distil (no) ↘ FAKE MEDIA BOND

The ideational targets are <u>underlined</u> and the attitude is in **bold.** The coupling of ideational and attitudinal meaning is expressed as: [ideation: ...../attitude: .....] to highlight the ideational target and its attitude sub-type. The social bonds (the broader values the community is aligning around) are written in SMALL CAPS and show in connection to the key affiliation strategy adopted to get people to rally around these bonds (with examples from the text in brackets). In this particular example, as the comment shares a DELIBERATE FIRE BOND (meaning the persona believes the false information about the Notre Dame Fire being deliberate, based on their negative evaluation of arson) and FAKE MEDIA BOND (meaning that they negatively evaluate the mainstream media as providing false information about the fire) it belongs to the Anti-Media persona. The label "Anti-Media" is based on the grounded theory analysis of social bonds in the dataset, where a significant number of comments negatively evaluated the mainstream media. The labelling reflects the main ideational target of these comments. To summarise, this is a qualitative method, relying on contextual knowledge in conjunction with discourse analysis in order to define social bonds and label textual personae.

4.3.6. Multimodal Analysis of Video Content

As YouTube is a multimodal platform, it is useful to also consider how the visual content of the videos is linked to information disorders. As explained in the data collection section, screenshots were taken of key frames in the videos. The multimodal analysis undertaken in this thesis considered the following dimensions of meaning:

- How do YouTube images compare to the meanings enacted in the transcripts?

- Do the images conform to genre conventions in terms of matching the genre identified by the textual analysis?

- How are images used as a form of 'visual evidence' to support deceptive or problematic communication claims?

The framework for analysing YouTube multimodal content was based on work by Bednarek (2014) where a metafunctional approach was taken to analysing each frame of a video. This involved exploring the following visual aspects of the video data:

- **Interpersonal**: This metafunction is focused on relationships and how people bond around deceptive ideas. The modality (Kress and Van Leeuwen, 2006) and legitimation (Van Leeuwen, 2007) frameworks were used to explore the interpersonal dimension of the videos.

- **Textual**: This metafunction is based around the organisation of a text. By considering YouTube videos as 'multimodal documents', work by Bateman (2008) on image, text and page flow, was used to investigate the textual dimension of the videos.

- **Ideational**: This metafunction is concerned with experience. By analysing the participants, action processes and circumstances based on work by Kress and Van Leeuwen (2006) the ideational dimension of the videos was explored.

This metafunctional perspective addresses the following issues relating to deception in the videos:

- How the image is mimicking trusted visual sources (modality and legitimation analysis)
- The organisation of the image (image, page and text flow)
- The most represented participants and circumstances in each image

A detailed breakdown of the specific analytical frameworks according to each of these metafunctional approaches will now be explained.

4.3.6.1. Visual Modality

In social semiotics, modality refers to "the truth value or credibility of (linguistically realised) statements about the world" Kress and Van Leeuwen (2006: 155). In other words, modality is concerned with how something is represented as true, not whether something is the absolute truth. Visual modality as a framework for this thesis is useful because it can tell us how these successfully deceptive or problematic YouTube videos construct believable content. Kress and van Leeuwen's visual modality framework suggests four coding orientations, that represent a set of principles for how texts are "coded by specific social groups, or within specific institutional contexts'" (Kress and Van Leeuwen, 2006: 165). These are the coding orientations:

- **Technological coding orientation**: this orientation represents how an image replicates a 'blueprint', typical to scientific and technological images. According to Kress and Van Leeuwen (2006), most images in this category are black and white, with colour therefore having a low modality.

- **Sensory coding orientation**: this orientation represents how an image heavily features sensory content, such as advertisements and artworks. In this sense, colour is affective and has a high modality.

- **Abstract coding orientation**: this orientation represents how an image has higher modality based on the more abstract and general it is, as opposed to concrete and naturalistic. To understand these images, one must understand particular academic, scientific, or artistic contexts.

- **Naturalistic coding orientation:** this orientation represents how an image doesn't require any particular knowledge to be understood. This can be seen as the default coding orientation in society, in the sense that this represents what people usually see without applying any specialist knowledge.

It should be noted that these coding orientations were developed for images rather than video content. Therefore, a challenge in this research is to adapt these coding orientations to multimodal YouTube content. Ravelli and Van Leeuwen (2018) have outlined some of these challenge regarding modality in the digital age, suggesting that "digital technology allows users to modify, at the press of a finger, the modality of images in ways which formerly would only have been available to experts" (Ravelli and Van Leeuwen, 2018: 288). With digital media, multiple coding orientations can now occur simultaneously rather than just a single coding orientation at a time.

For this thesis, these coding orientations were applied to each frame. A single coding orientation was applied to each frame or simultaneous coding orientations depending on the layout of each video frame. Adjustments were then made to the coding orientations based on patterns uncovered from the datasets and misalignments with the current coding orientations.

4.3.6.2. Legitimation

Legitimation (Van Leeuwen, 2007) is another framework that can be applied to understanding the visual content of videos and the legitimation strategies evident in the language of the transcripts and YouTube comments. By understanding how legitimation is established via visual content, this can show how deceptive content pretends to be credible.

The four main categories of legitimation developed by Van Leeuwen (2007: 92) include:

- **AUTHORISATION**: (De)Legitimation according to CUSTOM (traditions and conformity), AUTHORITY (authoritative people or authoritative objects) or COMMENDATION (experts or role models).

- **MORAL EVALUATION**: Legitimation of value systems via EVALUATION, ABSTRACTION, or COMPARISON.

- **RATIONALISATION**: Legitimation via institutionalised social action or the knowledge society has constructed, in terms of THEORETICAL LEGITIMATION (how knowledge is constructed in terms of the EXPERIENTIAL, SCIENTIFIC, DEFINITION, EXPLANATION or PREDICTION) or INSTRUMENTAL LEGITIMATION (focusing on institutionalised social action via the MEANS, GOALS or EFFECTS of actions).

- **MYTHOPOESIS**: Legitimation conveyed through narratives and future projections. For example, via MORAL TALES, CAUTIONARY TALES, SINGLE DETERMINATIONS (that represent stories in a straightforward way) or OVERDETERMINATIONS (that represent stories via INVERSION or SYMBOLISATION).

These four main categories also have sub-systems as shown in Figure 4-11.

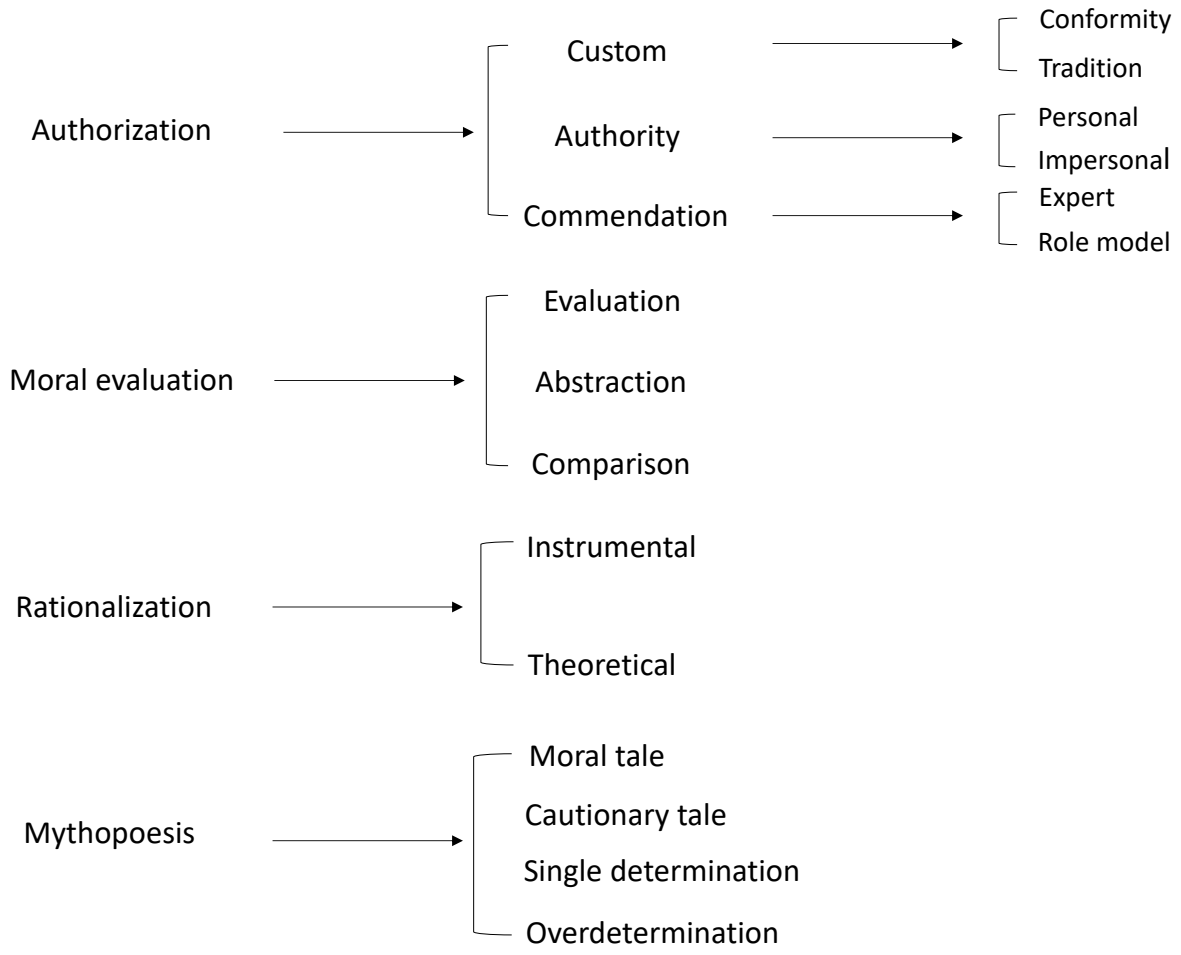


Figure 4-11 - Overview of the Legitimation Framework (van Leeuwen, 2007)

This system network of legitimation was considered in relation to how each frame establishes a sense of legitimation by its visual content, for example:

- Do authority figures establish legitimation?

- Can content like screenshots or newspaper clippings establish legitimation?

- By adding logos or promotional material to a frame does this establish legitimation?

In order to account for content typical to digital media such as screenshots, the legitimation framework was adjusted. The AUTHORITY sub-category is extended to consider further delicacy. The YouTubers in the dataset employed screenshots, video clips, and new articles as a form of evidence to legitimate their claims, and thus, we have developed the legitimation category of IMPERSONAL AUTHORITY as follows:

- **MARKETING**: Credibility established via company logos.

- **LAWS, RULES, AND REGULATIONS**: This sub-category incorporates the initial definition of impersonal authority by Van Leeuwen (2007), that is how references to laws, rules and regulations construct authority.

- **TECHNOLOGICAL**: Credibility established via technological means e.g., screenshots, online articles, and references to video links or google searches.

TECHNOLOGICAL AUTHORISATION is also refined to include:

- **ONLINE MEDIA**: Legitimation formed by referring to social media or online news articles as evidence.

- **TRADITIONAL MEDIA**: Legitimation formed by referring to newspapers and television as evidence.

- **TECHNOLOGIES**: Legitimation formed by referring to technologies (in the case of this study the only technologies referred to were aircrafts, e.g., drones and UFOs).

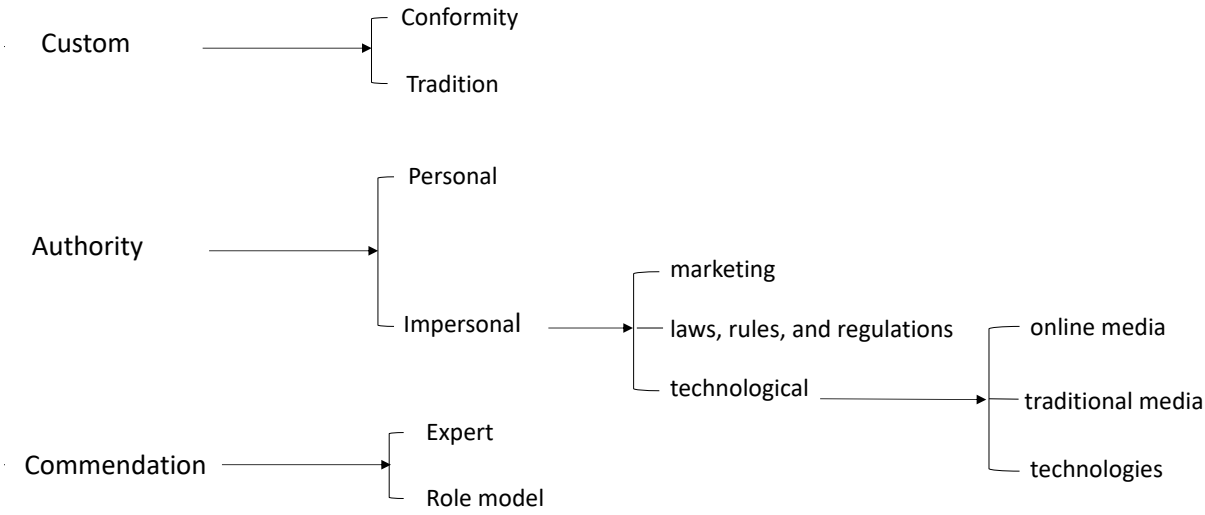


Figure 4-12 - Authorisation Framework with Added Sub-systems (adapted from van Leeuwen, 2007)

4.3.6.3. Image, Text and Page Flow

By considering Bateman's concepts of 'page-flow', 'text-flow' and 'image-flow', the way in which multiple elements work together in a single video frame can be realised (Bateman, 2008; Bateman, 2013). Originally, Bateman's GeM (Genre and Multimodality) framework was applied to static documents but has later been employed to 'non-static multimodal artefacts', encompassing audio-visual content (Bateman, 2013). This thesis codes each video frame according to the following:

- **Page Flow**: how one frame transitions to another frame e.g. does it stay on the same image, or does it switch to completely different content.

- **Text Flow**: how text works on each frame e.g. is there a banner of text continuously moving across the frame, or prominent static text on the frame.

- **Image Flow**: how images work on each frame e.g. are images continuously moving, or static for a longer period of time.

Taken altogether, this analysis reveals how YouTube videos can be seen as multimodal documents, with multiple elements existing on a single frame and continuously shifting. The results from this part of the analysis helped in determining what makes deceptive or problematic content believable, in terms of how various elements construct a sense of credibility. A summary of Bateman's concept of page-flow, text-flow, and image-flow, is provided in the following diagram:

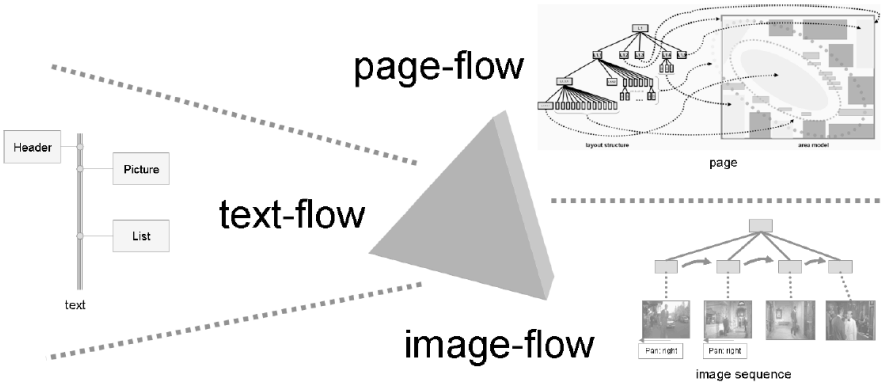


Figure 4-13 – Three Semiotic Modes Deployed Within Document Pages (from Bateman, 2008, p.175)

4.3.5.4. Visual Salience

The resource of visual salience was important in order to understand the relation of the visual elements in the screen shots with the social bonds realised by the values identified in the transcripts. From a multimodal perspective, visual salience refers to the elements in an image/video that are depicted as the worthiest of attention, formed, according to Kress and Van Leeuwen (2006: 202), through one or more of the following dimensions:

- ***Size***: larger elements rather than smaller elements
- ***Sharpness of focus:*** sharper rather than blurrier elements
- ***Tonal contrast***: elements with high tonal contrast
- ***Colour contrast***: strongly saturated elements

- ***Placement in the visual field***: elements in the centre or top of the image
- ***Perspective***: foregrounded elements
- ***Overlapping elements***: the element that is overlapping other elements
- ***Cultural significance***: the famous person in the image
- ***Personal significance***: an element that is more significant due to the viewer's personal experience

In this study, the visually salient element was primarily realised by the dimensions of size, placement in the visual field, overlapping elements, and cultural significance. Many of the salient visual elements observed in the video frames can also be interpreted as 'bonding icons', that is, symbols that embody particular values which people rally around (Stenglin, 2008; Tann, 2012; Zappavigna, 2014a; Zappavigna, 2014b).

4.3.6.4. Represented Participants, Action Processes, and Circumstances

This part of the analysis was derived from Bednarek's framework (2014) where a more simplified version of Kress and van Leeuwen's (2006) work is used to analyse ideational meanings. Bednarek's framework was selected because this framework has been previously applied to the frames of audio-visual content. Therefore, the following features were analysed based on Bednarek's framework (2014: 42):

- **Represented participants**: what participants are represented in each frame? This can include human and non-human participants.

- **Action processes**: what processes are the participants engaging in? The options here are either:
    - Dynamic or Non-Dynamic – is movement or no movement involved?
    - Gaze – is it a direct gaze towards the viewer, towards an object or no gaze at all?
    - Speech – is speech involved? Is the participant directly speaking or is it a voice-over?

- **Circumstances**: Where are the participants/processes located? A description of the setting is provided here.

Overall, this framework was useful for understanding "*who* is participating, *what* are they doing and *where* are they located?" (Bednarek, 2014: 42).

4.3.6.5. Summary of Multimodal Coding Strategy: Affiliation and Legitimation

The manual coding strategy used across the dataset of 30 videos in this thesis is outlined in Table 4-9. The aim was to reveal how the content of the videos is legitimated, evaluated, and how they function in the service of social bonding. The *(De)Legitimation* column in Table 4-9 provides the key categories of the legitimation framework, drawing on the analytical method previously detailed, alongside the key questions guiding the verbal and visual analysis, and the *Affiliation* column shows the key questions guiding the analysis, drawing on the analytical method detailed previously. The *Additional Multimodal Analysis* column highlights both further analyses that could be conducted (for example a phonological analysis on the verbal content) and the other multimodal analyes detailed in this chapter that complemented the affiliation and legitimation analyses (for example, a visual salience analysis delving deeper into how social bonds were made prominent in the video frames).

	(De)Legitimation	Affiliation	Additional Multimodal Analysis
Verbal content	**Authorization** – is authority (de)established via an authority figure or object? **Moral Evaluation** – does (de)legitimation occur via a value system? **Rationalization** – does (de)legitimation occur by reference to social actions and knowledge constructed by society? **Mythopoesis** – is (de)legitimation expressed by narratives that reward legitimate actions?	What is being evaluated? (the ideational target) How is it being evaluated? (what type of attitude) What key social bond is expressed?	Phonological analysis beyond the scope of this thesis
Visual content	**Authorization** – is authority (de)established via an authority figure or object? **Moral Evaluation** – does (de)legitimation occur via a value system? **Rationalization** – does (de)legitimation occur by reference to social actions and knowledge constructed by society? **Mythopoesis** – is (de)legitimation expressed by visual narratives that reward legitimate actions?	Is a social bond expressed in the visual content? (discourse iconography) Does this key visual bond align with what is being said (the language) or is it adding additional meaning?	What are the salient features of the selected video frame and how is it salient? (visual salience) What are the key participants, action processes, and circumstance? (visual salience) How is the selected video frame made to appear credible? (visual modality analysis accompanying legitimation analysis) How is the video frame organised? (image, page and text flow analysis)

Table 4-9 - Key Manual Coding Strategy for the Dataset

The multimodal affiliation and legitimation analyses for this study are presented according to conventions shown in Figure 4-14. The video frame is presented at the top of the diagram, with the salient elements in the frame annotated with a red rectangle. Salient visual semiosis contributing to the realisation of social bonds is shown via circular callouts, with corresponding visual legitimation strategies identified underneath the frame. A downward arrow symbol is used to indicate instantiated features. The transcript text (the YouTuber's verbiage when the screenshot appears) is presented in a speech bubble. Underneath the transcript are the key bonds realised by the ideation-attitude couplings in the transcript (ideation <u>underlined</u> and ATTITUDE in **bold**). Underneath this are any affiliation strategies, and legitimation strategies that are highlighted in the transcript text. In order to maintain ethical research standards, the faces of people who are not public figures have been anonymised with a black circle. This usually occurred in videos where the YouTuber speaks directly to the camera, as per the conventions of a vlog.

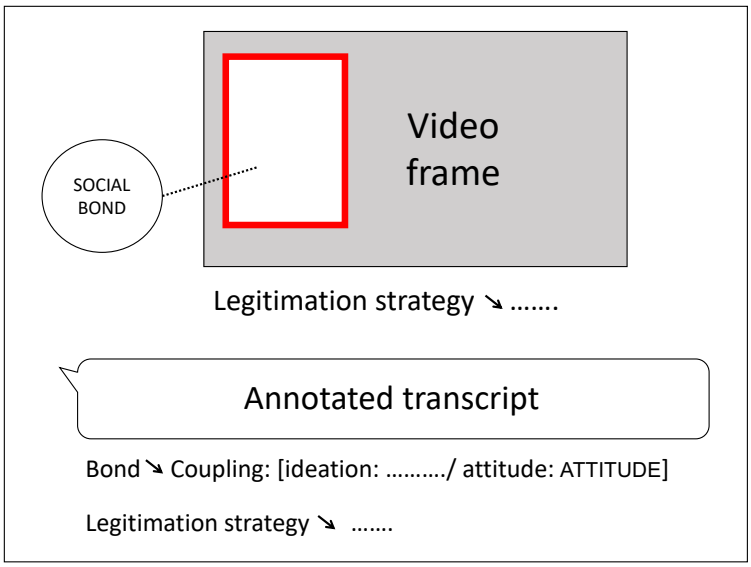


Figure 4-14 – Convention for Presenting the Multimodal Analysis of Affiliation and Legitimation

4.4. Conclusion

This chapter has explained the methodology and attendant methods applied in this thesis across the two case studies. Each case study follows a consistent method of data collection and analysis. The chapter has explained the details pertinent to each case study in terms of how specific videos were selected for analysis and the different analytical challenges that were encountered. Overall, the chapter has explained how key methods in SFL were used to answer each of the research questions. This SFL approach affords a metafunctional and multimodal perspective for understanding information disorders. It has enabled a focus on interpersonal meaning, a dimension of information disorders that has previously been neglected in studies that have tended to focus on whether information is true or false, rather than the bonds at stake that drive the proliferation of the false or problematic content.

4.4.1. Process Diagram

Because of the complexity of the sampling and analysis procedure that was needed in this these to explore the multimodal dataset, the following process diagram is provided as an overview (Figure *4-15*). It summarises the research process in terms of how the data was collected and analysed.

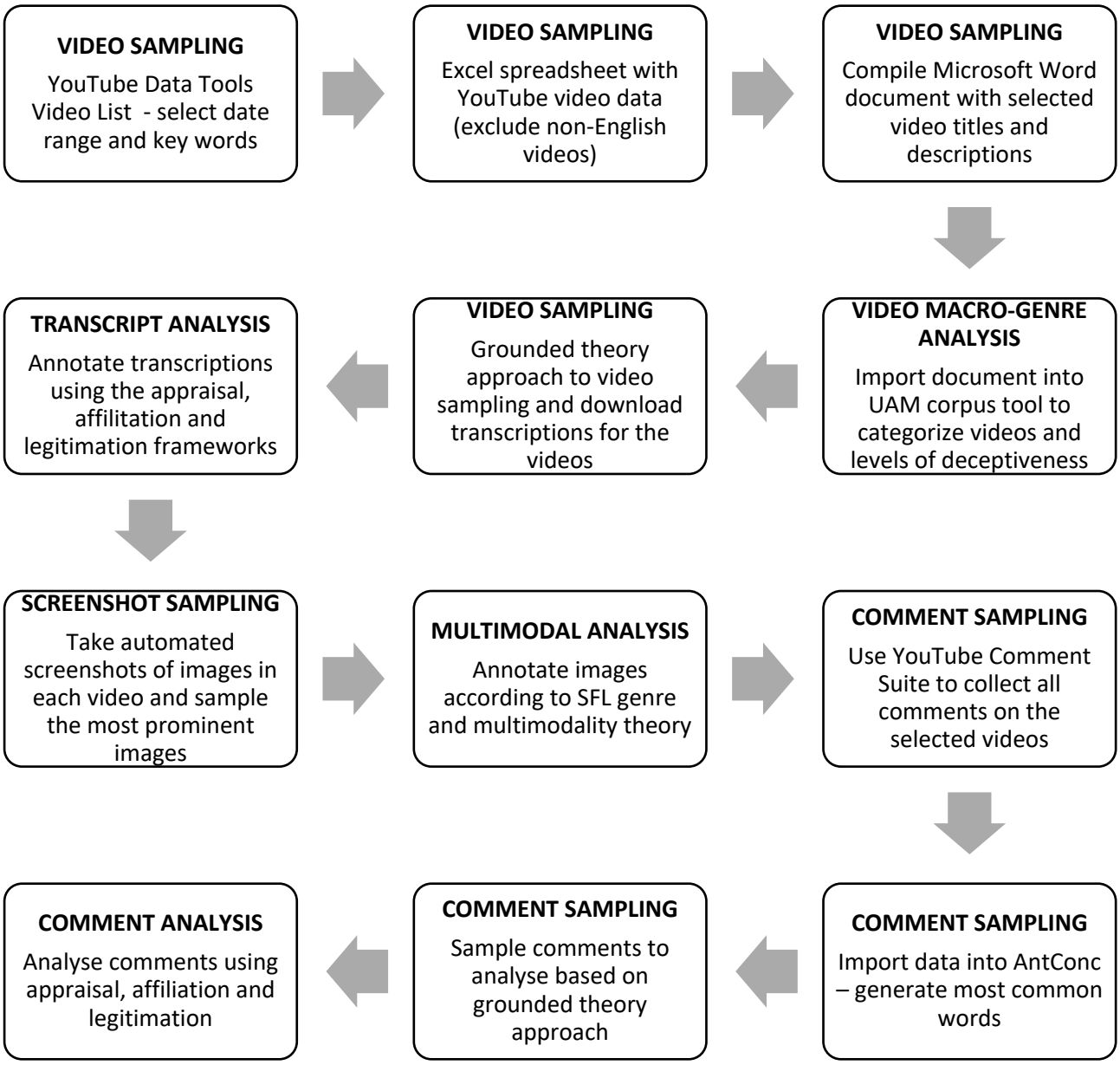


Figure 4-15 – Process Diagram of Methodology

5. CHAPTER FIVE THE MOMO CHALLENGE: NON-POLITICALLY MOTIVATED MORAL PANICS AND INTERNET HOAXES

5.1. Introduction

This chapter explores the first case study of this thesis – The Momo Challenge. The introduction will provide some further context on the Momo Challenge and how it became a viral phenomenon online. The content analysis will explain the results from the analysis conducted on the metadata of the videos and the creation of 'macro-genres' in order to understand the broader differences among the video dataset. The next section, focuses on the transcript and visual analysis of the videos, taking into consideration the role of affiliation and legitimation in the shaping of deceptive discourses and moral panics within the videos, and the role of technological evidence to support claims. The main findings will then be summarised, and detailed illustrative examples representing the two main macro-genres from the dataset will be provided in order to further understand the genre stages of each video. The social bonds at stake in information disorders and moral panics are explored in terms of both how they are constructed visually and in language, as well as how these bonds are legitimated in the discourse.

5.1.1. Background Context: The Momo Challenge

This section will provide some background on the history of the Momo Challenge. The Momo Challenge reached its peak as a viral internet sensation in February 2019 (refer to *The Methodology Chapter* for more details). During this time, various news outlets and concerned parents were concerned about children receiving threatening messages on WhatsApp and being shown images of a frightening figure. Other variations of this

story also reported that footage of the frightening figure was spliced into YouTube Kids videos such as 'Peppa Pig' videos[14].

The Momo challenge was incorrectly linked to the suicide of children from around the world. No evidence has been discovered or statements from law enforcement agencies that directly links the Momo Challenge to the death of any children[15]. The Momo Challenge shares many similarities to another non-existent social media challenge called the 'Blue Whale Challenge' that was first reported on by the Russian newspaper *Novaya Gazeta*[16]. In addition, it has been linked to a 'creepypasta' story on Reddit – where users share fictious horror stories with one another, that they have copied and pasted from other online sites[17]. Nonetheless, due to the viral spread and meme-like status of the Momo challenge, it cannot be conclusively linked to one original source. The Momo Challenge has now been discredited by most major news outlets and YouTube has stated that it did not discover any Momo Challenge spliced videos[18]. Now the Momo Challenge is more commonly referred to as an 'internet hoax' and an 'internet urban legend' [19].

---

[14] For example, see this article from *The Sun*: https://www.thesun.co.uk/news/8515907/momo-challenge-fortnite-youtube-instagram-self-harm/

[15] News articles discrediting the momo challenge include: *The Guardian* https://www.theguardian.com/technology/2019/feb/28/viral-momo-challenge-is-a-malicious-hoax-say-charities ; *The Atlantic* https://www.theatlantic.com/technology/archive/2019/02/momo-challenge-hoax/583825/ ; and fact-checking site *Snopes* https://www.snopes.com/news/2019/02/26/momo-challenge-suicide-game/

[16] Discussions about the Blue Whale Challenge are featured on *Know Your Meme*: https://knowyourmeme.com/memes/blue-whale-challenge and *Wikipedia*: https://en.wikipedia.org/wiki/Blue_Whale_Challenge

[17] For example, the website *Know Your Meme* traces the origins of momo: https://knowyourmeme.com/memes/momo-challenge

[18] YouTube's statement can be accessed here: https://support.google.com/youtube/thread/1917881?hl=en

[19] Discussion of experts falling for the Momo challenge internet hoax: https://theconversation.com/momo-challenge-shows-how-even-experts-are-falling-for-digital-hoaxes-112782

Another site that captures evolving pop culture trends, knowyourmeme.com, outlines a similar timeline of the Momo Challenge. This site in particular, traces Reddit posts, YouTube videos, memes posted on social media, and tweets that went viral. Nonetheless, whilst the Wikipedia and 'knowyourmeme' pages refer to key sources that spread the Momo Challenge, its exact origins are unknown. Similar to memes, the Momo Challenge cannot be definitively traced, and reached a level of virality, where it was continuously replicated and became part of 'copycat' attacks, as indicated by these comments from Wikipedians:

> there are only trolls and predators that use it as a masquerade to fulfill their desires using a panic

> hoax challenges 'becoming real' through unironic participation

People may have created accounts on WhatsApp pretending to be Momo once the Momo Challenge had reached virality, however, the stories of suicides and self-harm linked to the Momo Challenge are false. In addition, as this chapter will show, many YouTubers continued spreading false information about the Momo Challenge, despite knowing that it was a hoax, in order to achieve viral videos. Thus, the Momo Challenge as a case study offers many layers of information disorder in terms of misinformation, disinformation, and click-bait.

5.2. Content Analysis of Video Metadata

This section outlines the initial content analysis conducted on the metadata of YouTube videos about the Momo Challenge. The process for gathering these videos is outlined in *The Methodology Chapter*. The 15 videos selected for closer analysis each represent a potential macro-genre that was identified in the larger analysis of the 195 videos initially collected.

5.2.1. **Video Macro-Genres**

Five potential macro-genres were identified by the initial content analysis. They are defined as potential macro-genres since they can involve multiple elemental genres, but 'potential' in the sense that these macro-genres would only be confirmed after the more detailed analysis as outlined in *The Methodology Chapter*. These macro-genres were:

- **News reporting** – Videos that clearly identified in the title or description that they were from an established news station. In this case, the news station has a presence beyond YouTube and features on a TV channel. This could be identified by the names of the channels, for example:

  Denver7 - The Denver Channel
  WPTV News | West Palm Beach Florida
  Inside Edition

- **Commentary** – Videos that provided opinions about the Momo Challenge but were not directed towards a particular audience. In the descriptions of these videos, viewers were encouraged to listen to the videos in order to know more about the Momo Challenge:

  Exploring The Momo Situation
  If you want to stay in the know of more pop culture happenings be sure to hit that subscribe button.
  ...to have more information about this new trending suicidal apps better watch and share this video to everyone.

- **Educational** – Videos that outlined in the description or title that the main aim was to educate people. Often, the target audience was indicated as parents. The video titles and descriptions also focused on understanding 'how' the Momo challenge works, for instance:

Take Notice Parents the momo challenge and Peppa pig videos are not fake. They are real.
How the Momo Challenge uses Psychology to Kill
The Momo Challenge & Creepy Peppa Pig Videos? (PARENTS WATCH THIS)

- **Entertainment** – Videos that focused on entertaining audiences about the Momo Challenge by either trying to call Momo or reacting to videos about Momo:

While trying to debunk the Momo challenge we actually made contact through whatsapp with Momo and got it on camera.
"THE MOMO CHALLENGE" Peppa Pig ORIGINAL VIDEO REACTION
YOUTUBERS REACT TO MOMO (Scary Meme or Hoax?)

- **Advertising** – Videos that appeared to be 'click-bait' based on reading the title and video description. Titles were often written in all caps and attracted attention to the video by stating that something would be 'exposed'.

MOMO CHALLENGE EXPOSED | DON'T DO IT | DO NOT CALL PERIOD | WHATSAPP GAME HOAX WARNING
*EXPOSING THE MOMO CHALLENGE *REAL VIDEO EVIDENCES**
ITS REALL!!!! MOMO CHALLENGE REACTION!! (IM SCARED)

Based on the descriptions of each video, each video was also categorised according to its level of deceptiveness:

- **Satire** – Video spreads false information but belongs to the entertainment genre.

- **Unintentional misinformation** – The video spreads false information but it does not appear that the video creator is intending to deceive someone, rather the video creator believes that their content is correct.

- **Intentional disinformation** – The video was intentionally deceptive e.g. uses click-bait titles and/or is trying to sell a product/brand rather than reporting correctly on the Momo Challenge.

- **Unclear** – Based on the title and description of the video, it was not possible to categorise the video. This could be due to short non-descriptive titles or a blank video description section.

It is important to highlight that this macro-genre analysis was solely based on the title and description of each video. The content of the video was not watched prior to this analysis. When searching for videos on YouTube, usually the individual only first sees the title and description of a video, and so the aim of this initial analysis was to categorise videos similar to this process. Thus, these potential macro-genres are only confirmed in the detailed analysis undertaken in the following section.

5.3. Analysis of YouTube Videos: Affiliation and Multimodal Legitimation Strategies

This section outlines the results of the analysis conducted on the transcripts and visual content of the 15 Momo Challenge videos primarily using multimodal discourse analysis methods in affiliation and legitimation. For the initial analysis, the three metafunctions were also considered: interpersonal (visual modality), textual (flow and salience) and ideational (represented participants, action processes, and circumstances). Nonetheless, the main focus of this analysis is on the interpersonal metafunction, conducting a hybrid affiliation and legitimation analysis.

The research questions informing the video transcript and visual content analysis were:

- What key social bonds are reflected in the transcripts and how are these bonds legitimated?
- What are the bonding icons reflected in the visual content of the videos?

- What visual resources are used to legitimate the social bonds in the transcripts?
- Are there any key linguistic patterns that suggest particular strategies common to the videos? Can any generalisations be made regarding the key linguistic strategies of certain macro-genres of videos?

This transcript and visual analysis will delve deeper into issues of deceptive evidence and reveal how using an SFL-based affiliation and legitimation analysis can highlight inconsistencies in attitudes towards evidence across the texts. Therefore, this section will explain the key legitimation strategies discovered across the videos as a starting point, and then detail within each of these strategies, the attitudinal language, social bonds, and legitimation strategies of the presenters embodying these strategies. This is exemplified by samples from the transcripts and video frames.

5.3.1. **News Stories and Technological Authority**

This section will provide an analysis of the role of new stories as technological authority, in other words, how new stories are used as evidence to legitimate social bonds mainly around Momo being a real and dangerous threat. Each YouTube video expressed a reliance on news stories to rhetorically assert the newsworthiness (Bednarek and Caple, 2017) of the Momo Challenge regardless of the macro-genre they belonged to. In terms of macro-genre structure, videos that belonged to the news-reporting macro-genre incorporated legitimation strategies that relied on the voices of 'experts' to construct an interesting news story (as will be explored in Section 5.4.). Videos from other macro-genres (like entertainment) also incorporated legitimation strategies that involved using news stories to introduce the subject matter of the Momo Challenge, before the YouTuber would share their own warnings and opinions about the Momo Challenge.

The numerous references to news stations at the start of videos in the Momo Challenge dataset legitimises the Momo Challenge as a worthy news story. As previously explored in *The Methodology Chapter*, the majority of videos in the Momo Challenge dataset did not belong to the news macro-genre, rather these videos were typically in a vlog format, providing commentary or entertainment regarding the Momo Challenge. Despite this,

these videos still reference new stations frequently in their commentary, by showing short clips of news stories and referring to its current virality. In the following example (Figure 5-1), the reference to Momo being on "*tons of news stations*" legitimatises it as real and not an internet hoax.

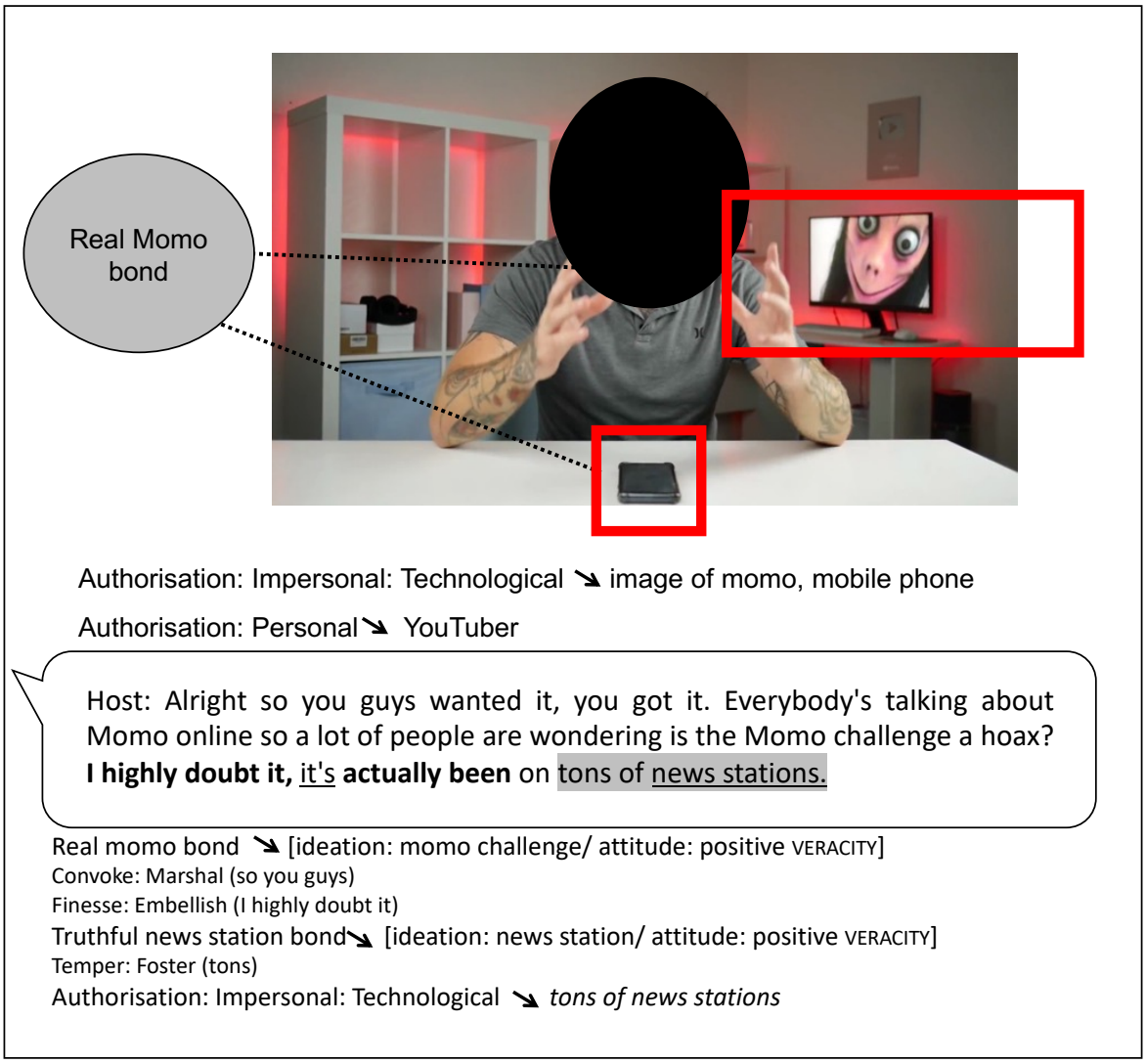


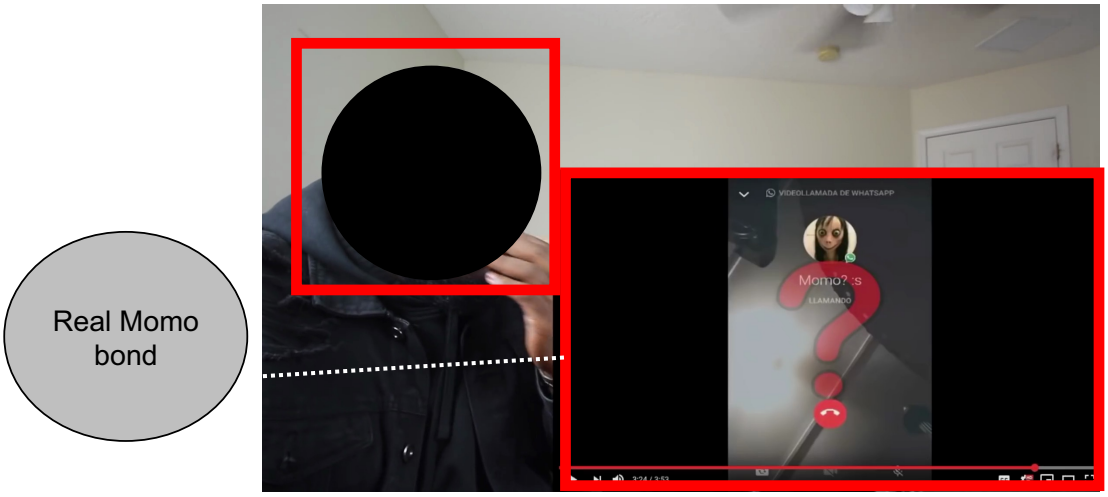
Figure 5-1 - Example 1 of News Stories as Technological Authority

Firstly, the YouTuber uses a convoking affiliation resource "*you guys*" to address their imagined audience, before evaluating Momo as a real entity with positive veracity ("*…is the Momo Challenge a hoax? I highly doubt it*") not an internet hoax. "*News stations*" are an entity which invokes positive evaluation, therefore seen as truthful. This reference to "*news stations*" acts as the legitimation strategy used to ascertain that Momo is real. Visually, the "real Momo bond" is evident by the placement of the mobile phone near

the YouTuber. This is frequently referred to visually and verbally throughout the video as evidence of Momo existing, as the YouTuber would then hold the phone up to the camera to show a series of WhatsApp messages supposedly received from Momo. An image of Momo in the video acts as a 'bonding icon', in other words, a symbol that rallies people in their fascination of Momo as a frightening and mysterious figure. The visual impact of Momo is a crucial element in persuading the audience of Momo's significance. Hence, there are two key visual legitimation strategies occurring in this video. Firstly, the personal authority of the YouTuber, shown as a salient figure in the centre of the frame that is directly speaking to the camera and sharing their thoughts. Secondly, the mobile phone and image of Momo are representative of TECHNOLOGICAL AUTHORITY because the mobile phone is also centred in the frame and later shown as evidence of Momo existing due to its text messages, and the image of Momo features throughout the video in the background but nonetheless is also salient due to its cultural significance (a viral meme). This example highlights the multiple layers of legitimation constructed in language and visually.

In another example (Figure 5-2), reference to the news is also used to assert the validity of the Momo challenge. In this example, the YouTuber similar to the previous example addresses their imagined audience via convoking affiliation resources "*you know what I'm saying*" and states that their parents heard about the Momo Challenge on the news, with an invoked evaluation of the parents' reaction to Momo as NEGATIVE INSECURITY due to their uncertainty. Momo also has an invoked evaluation of Momo as an entity that is dangerous and that people like the YouTuber's parents are frightened about. "*The news*", a repeated phrase in this transcript, is an example of TECHNOLOGICAL AUTHORITY because it strengthens the dangerous Momo and frightened parents' bonds promoted in the transcript excerpt. Visually, the real Momo bond is also promoted via TECHNOLOGICAL AUTHORITY in the form of the YouTube clip in the bottom right-hand corner of the screen. This was a common structure with the YouTuber vlog format, consisting of simultaneous screens: the YouTuber speaking and a YouTube clip as the added evidence for the YouTuber's claims, in other words, the TECHNOLOGICAL AUTHORITY. As this example shows, TECHNOLOGICAL AUTHORITY is constructed via multiple layers, visually as

in the representation of additional YouTube clips, and through language by repeated reference to media forms such as the "*news*".



Authorisation: Impersonal: Technological ↘ video clip of Momo

Authorisation: Personal ↘ YouTuber

You know what I'm saying, parents **wouldn't be this**, what is on the news. Bra my mom **seeing it a minute ago in the news**. Momo challenge and this is oh this been **going on for like a year and it's trending again** in 2019.

Frightened parents bond ↘ [ideation: parents/ attitude: negative INSECURITY]
Convoke: Marshal (you know what I'm saying, bra)
Dangerous Momo bond ↘ [ideation: it (Momo)/ attitude: negative PROPRIETY]
Finesse: Distil (wouldn't be this)
Authorisation: Impersonal: Technological ↘ *the news*

Figure 5-2 – Example 2 of News Stories as Technological Authority

A recurrent legitimation strategy throughout the video dataset was based on the newsworthiness of the Momo Challenge. This newsworthiness was constructed via reference to the broadcast news, ranging from videos that were actual news broadcasts of the Momo Challenge, to videos that referenced these news broadcasts in order to strengthen the bond that Momo is real and significant. As the two examples discussed show, these social bonds were also strengthened via convoking affiliation resources, that is, resources that directly addressed listeners (e.g. *you guys, you know what I'm saying*) thus establishing a sense of community who are interested in the Momo

Challenge, particularly when it is emphasised as newsworthy or trending. The reference to news stories is an example of TECHNOLOGICAL AUTHORITY because technology itself (in the form of naming the news or including a YouTube clip) acts as the form of authority to legitimate the social bonds that the YouTuber propagates. As shown, the repeated reference to news stories, has an important role across various macro-genres in the Momo Challenge dataset.

5.3.2. **Experts and Authorisation**

Despite the conspiratorial or sensationalist nature of the Momo Challenge videos, reference to experts was still a legitimation strategy in the dataset. In particular, reference to experts was common in the news reporting macro-genre of Momo Challenge videos. With the news reporting macro-genre, after providing some context about the Momo Challenge, interviews were conducted with an "*expert*" – ranging from a principal at a high school, a psychologist, or a technology or cyber-security expert. Nonetheless, as this analysis will show, the reference to experts was misaligned, in other words, experts were used to legitimate social bonds, that they did not actually have any expertise in.

As a legitimation strategy both visually and verbally, experts were strategically placed in the news reporting macro-genre videos. An example of how an expert was introduced in a video is shown in Figure 5.3.

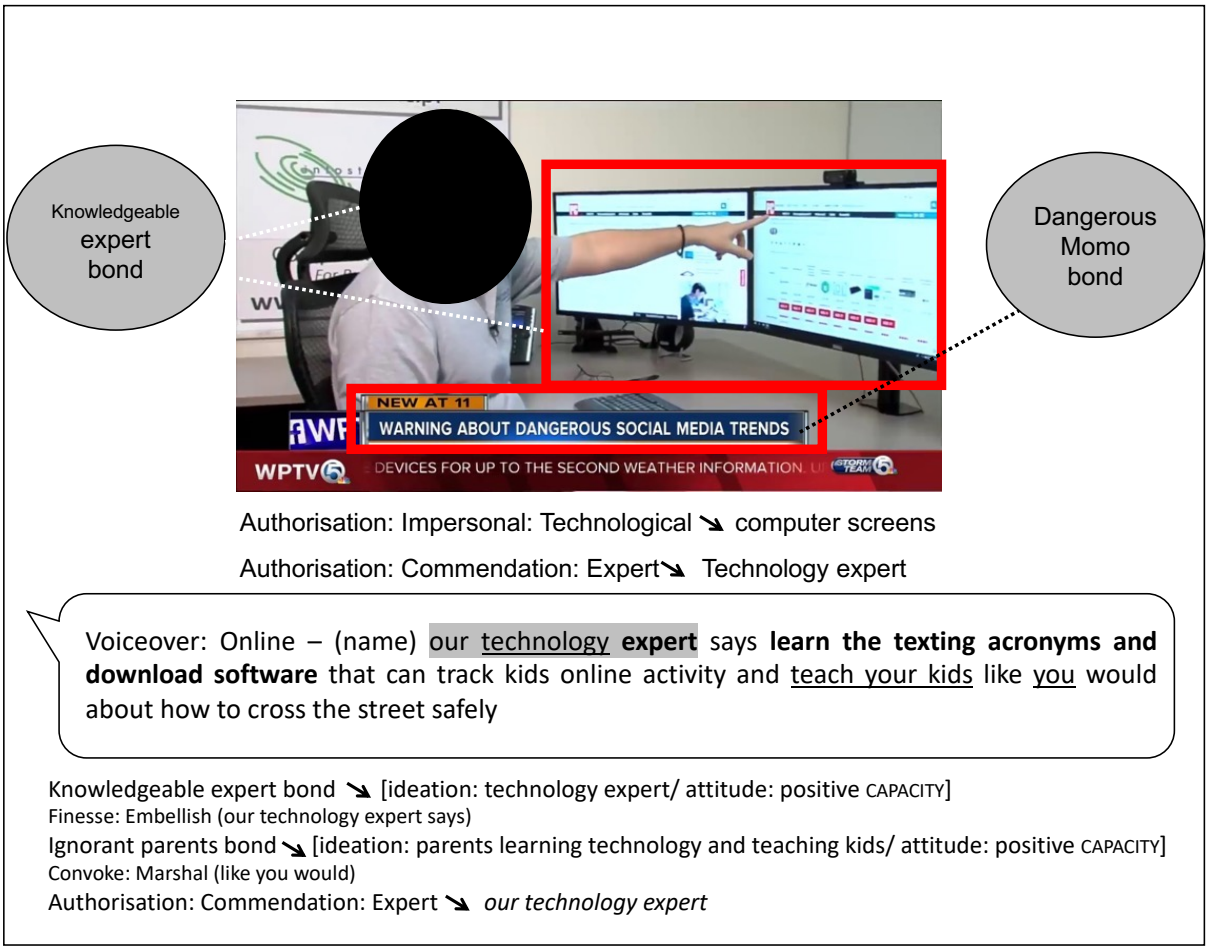


Figure 5-3 – Example 1 of Experts and Authorisation

In this example, a technology expert states that parents should learn texting acronyms and download software that can track kids' online activity, analogising this to how one would teach their kids "*about how to across the street safety*". Despite these experts being introduced as knowledgeable in the field, often their knowledge did not translate to the actual concerns of the Momo Challenge. For example, the technology expert in this example was more focused on "*texting acronyms*" than the details of the Momo Challenge. Momo is not the ideational target, instead it is the notion that parents should learn more about technology and teach their kids about technology.

In terms of legitimation strategies, the reference to a "*technology expert*" takes precedence in the transcript excerpt, legitimising the social bond of parents being ignorant about their child's behaviour online. Visually, this is complimented by the footage of the expert and the computer screens as a form of legitimation, serving the

notion that the expert is knowledgeable. The dynamic banner across the bottom of the video also forms a "dangerous Momo bond" as it negatively evaluates "*social media trends*". Nonetheless, similar to the transcript excerpt, there is a disconnection between what is shown on the banner versus the content of the computer screen (a webpage about internet acronyms). This misalignment regarding social bonds highlights how the expertise of the expert is not actually relevant to the Momo Challenge. Thus, videos in the news reporting macro-genre, often moved away from the original ideational target of the Momo Challenge, instead discussing issues around technology, and moving away from discussion about the Momo Challenge as a verifiable incident. This is a distraction from questions regarding the Momo Challenge's veracity, which should be the main subject in focus.

Expertise as a legitimation strategy also extended to visual references that acknowledged law enforcement agencies. This expertise in the news reporting macro-genre was commonly shown through emulated screenshots. Emulated screenshots encompass screenshots that were not actually taken by someone but are instead imitations of screenshots that have been heavily edited or created entirely from scratch. These functioned to provide TECHNOLOGICAL AUTHORITY in scenarios where the YouTuber was unable to access or to publish primary source texts. *Figure 5-4* consists of an emulation of the layout of a tweet in the background. The foreground is initially presented as a blank tweet template that was gradually populated with images of changing screenshots and PowerPoint-like animations of changing text. The purposes of these emulated screenshots are not explicit. It could be speculated that they are perhaps trying to recreate an event that can't be documented as the originals were no longer available or there were copyright issues relating to the news broadcast. The emulated screenshot may also have been used for aesthetic reasons in order to highlight 'internet culture' in the news story.

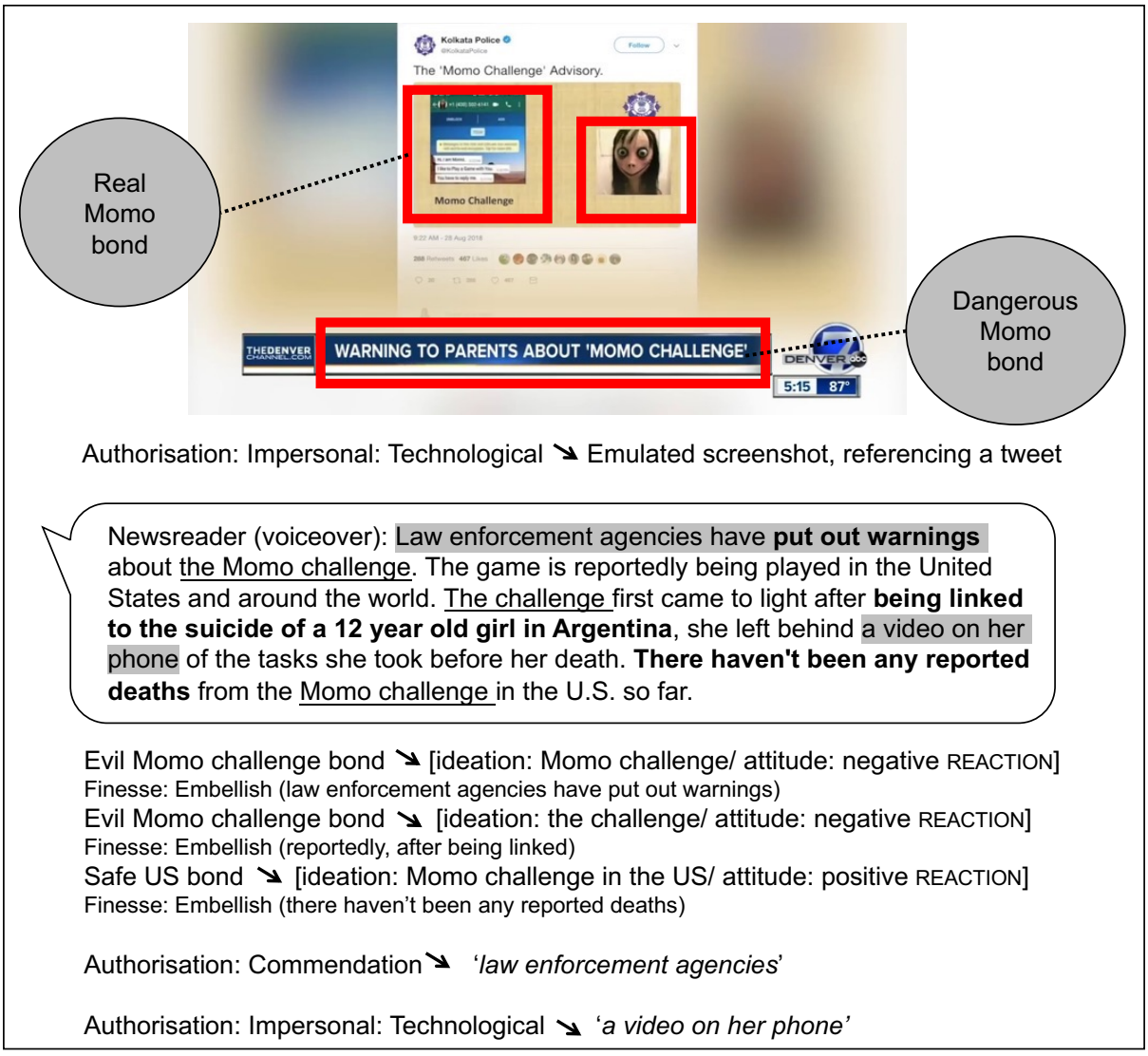


Figure 5-4 – Example 2 of Experts and Authorisation

The video voice-overs provide further insight into the kinds of meanings being made with these emulated screenshots. The Momo challenge is negatively evaluated, realising an 'Evil Momo challenge bond'. This matches the visual content with the salient image of Momo signalling danger. The reference to leaving behind a '*video on her phone*' also aligns with the emulated screenshot trying to recreate, as a form of TECHNOLOGICAL AUTHORITY a sense of evidence. However, there are also some additional meanings formed from the video voice-over. Firstly, the Momo challenge in the US is the target of invoked positive REACTION (because no deaths have currently been reported in the US) forming a 'Safe US bond'. Additionally, there is also some COMMENDATION, with the reference to law enforcement agencies legitimating the claim that Momo is a danger.

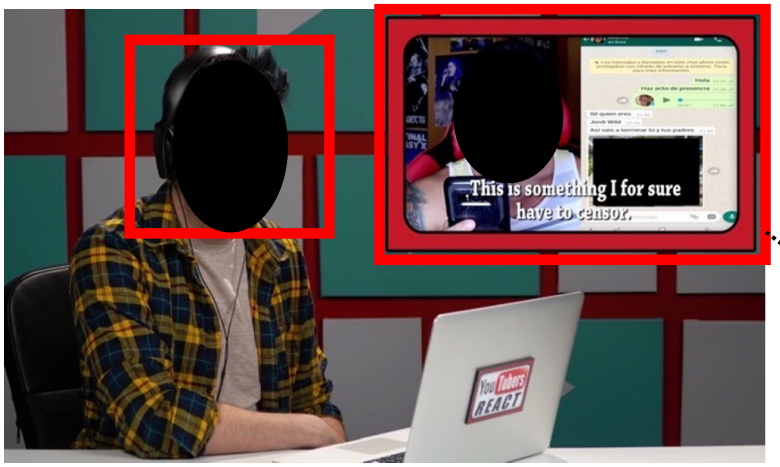
Overall, the function of the emulated screenshot appears to be to illustrate findings that cannot be shown as positive proof.

Overall, the depiction of experts was prominent across the entire video dataset. In the news reporting macro-genre, videos followed a formulaic news structure, incorporating voices of supposed experts, despite these experts not actually focusing on Momo as the main ideational target. In other macro-genres (as discussed in the previous section), clips and references to news stories and experts in the dialogue were used as evidence to support claims about the veracity of the Momo Challenge. Following on from these observations, the next section will discuss the wider role of intertextuality in the YouTube videos.

5.3.3. **Intertextuality and Technological Authority**

Each of the videos in the dataset relied on intertextual references as a legitimation strategy to construct a compelling account of the Momo Challenge. This section will discuss how intertextual references, typically represented as YouTube clips embedded in the corner of the vlogger's screen or screenshots, act as a form of TECHNOLOGICAL AUTHORITY. This TECHNOLOGICAL AUTHORITY supports the key social bonds propagated by the YouTuber, persuading audiences that the Momo Challenge is real.

Split screen screenshots functioned as a form of intertextual reference and had a legitimating function in the Momo Challenge dataset. This dataset also contained examples of an inverse split screen structure, with the YouTuber in the dominant visual position and the screenshot of a supposed interaction with Momo in the top right corner:



Truthful Information bond

Authorisation: Impersonal: Technological ↘ WhatsApp screenshot, laptop

Authorisation: Personal: Individual ↘ YouTuber

(Narrator) If you can get Momo to interact with you, you'll most likely be met with **insults, implications that this person knows your personal information**, and perhaps most notably, **disturbing** images that I would not be able to show you here.

Evil Momo bond ↘ [ideation: Momo/ attitude: negative PROPRIETY]

Convoke: Marshal (if you can get, show you)

Disturbing Images bond ↘ [ideation: images/ attitude: negative REACTION]

Temper: Modulate (most likely, most notably)

Mythopoesis ↘ *'if you can get Momo to interact with you'*

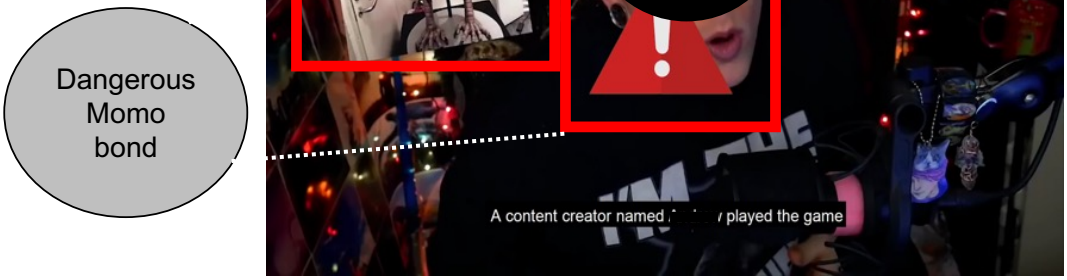
Authorization: Impersonal: Technological ↘ *'images'*

Figure 5-5 – Example 1 of Intertextuality and Technological Authority

Whilst the YouTuber and their reaction to the content in the screenshot takes prominence visually (in terms of the amount of space taken), the screenshot remains a bonding icon, with its visual presence promoting a 'Truthful Information bond', as screenshots have a cultural value as "receipts" of a past event (Jaynes, 2020). In the voice-over, Momo is the target of negative PROPRIETY for insulting and manipulating people, and images (most likely of Momo) are the target of negative REACTION. Even though the YouTuber claims these images cannot be shown to this audience, pointing to their existence itself serves as a form of TECHNOLOGICAL AUTHORITY, with the self-censorship heightening their supposed significance. The additional legitimation strategy of

MYTHOPOESIS is also realised when the YouTuber begins a recount of what will occur if '*you can get Momo to interact with you*'. Based on the analysis of the split screen throughout the video dataset, it appears that the split screen has two main functions: it either privileges the screenshot as an essential piece of information or privileges the YouTuber speaking as a legitimate voice of knowledge in the video with the screenshot adopting a more auxiliary evidentiary role.

Intertextuality as a legitimation strategy was intertwined throughout the videos in the dataset. For example, throughout the dataset, when YouTube presenters were talking, they alluded to other YouTube videos about the Momo Challenge, using this intertextuality as evidence for their claims. This can be seen in the following example (Figure 5-6):



Authorisation: Impersonal: Technological ↘ image of Momo, added YouTube footage

Authorisation: Personal ↘ YouTuber

A content creator and gamer named [name] played the game to show how **dangerous** it can be. He narrates what happened in his channel called [title] according to him Momo will call and message the user, she will **torment** and **encourage them to harm themselves**.

Dangerous Momo game bond ↘ [ideation: Momo game/ attitude: negative REACTION]
Finesse: Embellish ([name] played the game)
Temper: Modulate (how dangerous)
Dangerous Momo bond ↘ [ideation: Momo/ attitude: negative PROPRIETY]
Finesse: Embellish (according to him)

Authorisation: Personal ↘ *content creator and gamer named [name]*
Authorisation: Impersonal: Technological ↘ *channel called [title]*

Figure 5-6 - Example 2 of Intertextuality and Technological Authority

In this example, reference to another YouTube channel takes precedence visually. The example shows found footage of the YouTube channel clip, with the YouTuber then overlaying an image of Momo and their own animated red exclamation icon in the centre of the clip. This found footage, is overlayed with the YouTuber's own narration and icons. In particular, the icons legitimate a dangerous Momo bond, with the red exclamation icon centred in the frame and animated so that it becomes larger and larger as the YouTuber speaks, creating a sense of urgency regarding the dangers of the Momo Challenge. These legitimation strategies are also replicated in the transcript text, with the negative evaluations of Momo establishing a dangerous Momo bond legitimised by

reference to the YouTuber "*(name)*" (PERSONAL AUTHORISATION) and his channel (title) (TECHNOLOGICAL AUTHORITY). This example shows another strategy for how intertextual references are weaved into YouTube videos.

As shown by these examples, intertextuality among YouTube videos and discussion of the dismissal or use of evidence to back-up claims, acted as prominent legitimation strategies and as a way to rally around commonly viewed videos throughout the dataset. The particular dataset for analysis contained videos that all made reference to at least one other video in the dataset. As explored in *The Methodology Chapter*, the videos in the dataset were the most popular ones on the topic of the Momo Challenge and so this intertextuality also highlights how these videos promote each other and increase a video's particular reach, and assume that the viewer may have already watched or will watch other videos on the Momo Challenge. The most common way of representing intertextuality was via split screens, with these split screens showing other YouTube clips and screenshots.

5.3.4. **Evidence and Technological Authority**

Despite the conspiratorial nature of the Momo Challenge videos, there was significant discussion about the reliability of evidence as a legitimation strategy to strengthen the YouTuber's own sense of authority. In particular, screenshots of social media posts served a legitimation function, drawing on TECHNOLOGICAL AUTHORITY as evidence that the Momo Challenge is a real, ongoing, and tangible threat. This section will explore specific examples of screenshots legitimating the social bond that Momo is real.

In the dataset, there were examples of dynamic screen recordings of social media feeds used for legitimation. An example is shown in the following:

Figure 5-7 – Example 1 of Evidence and Technological Authority

Visually, what is striking about these screen recordings is the salient image of 'Momo', who is a recognisably threatening figure associated with negative emotions such as fear. The voice-over, negatively APPRECIATES the challenge as '*nothing new*' ([ideation: Momo Challenge / attitude: negative VALUATION],), invoking negative judgement of people who do not recognise its ongoing history. This is also part of enacting a 'Real Momo' bond that is shared by people spreading the misinformation. The legitimation strategies employed here again employ TECHNOLOGICAL AUTHORITY, with the YouTuber referring to practices of using technology ('*scrolled through Facebook, Instagram*'). The visual salience of the Momo image in the screen recording strengthens the overall bond that the Momo Challenge is a real phenomenon.

Screenshots of online articles were used as evidence for the claim that the Momo Challenge has resulted in the death of children. These articles were often from local media or non-mainstream media outlets, or were cropped in such a way as to only highlight the title without further context. In *Figure 5-8* Momo is again the salient image drawing the viewer's attention. The voice-over realises a 'Dangerous Momo bond' by negatively VALUING the challenge, linking it to the suicide of a 12-year-old girl and 16-year-old boy. The source entity, '*local media reported*', evokes IMPERSONAL AUTHORITY. The lack of detail, in terms of the specific local media organisation, mirrors the scant detail in the screenshot of the online article. This is an example of the way that these sorts of screenshots can deceptively distort news stories.

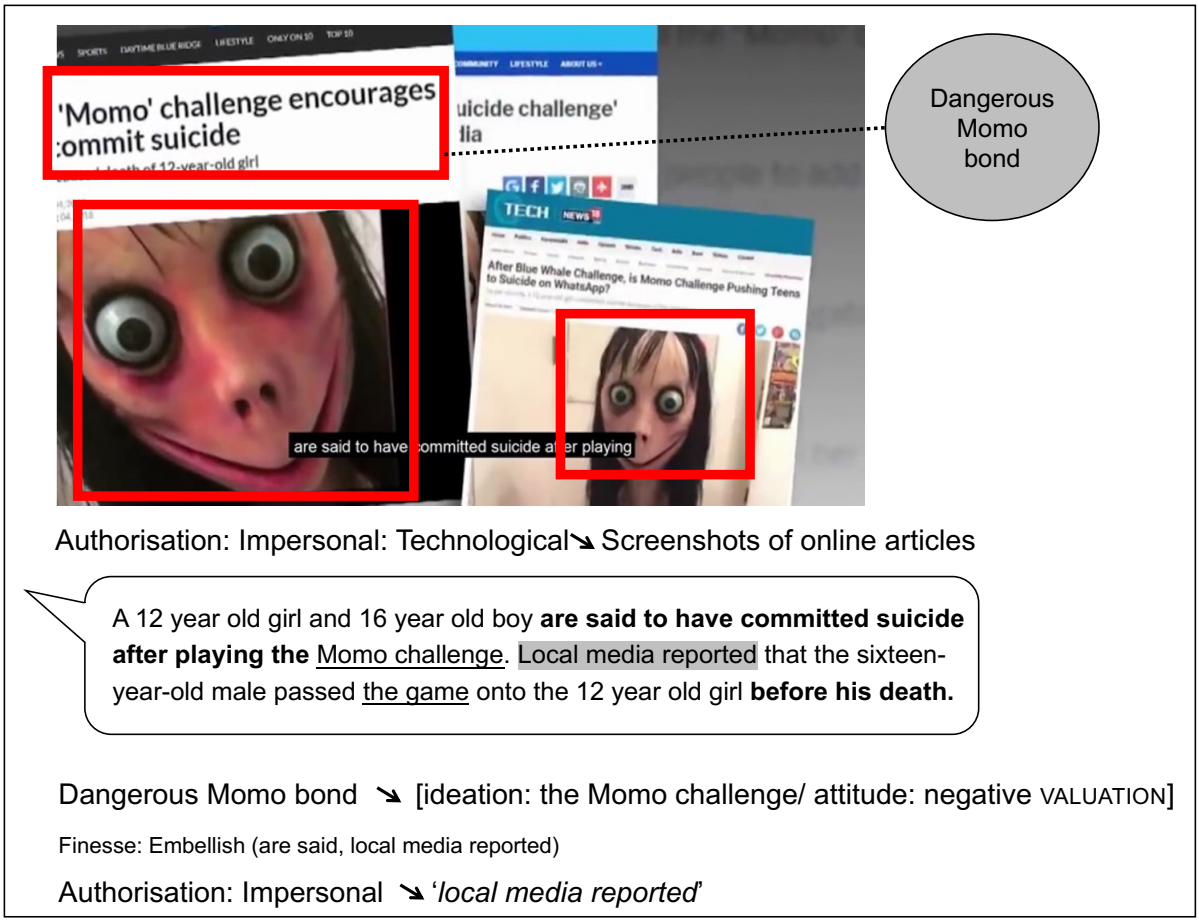
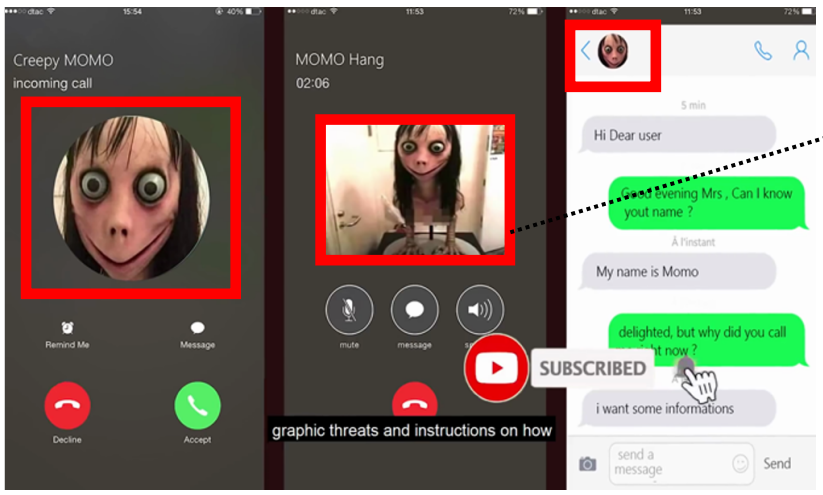


Figure 5-8 – Example 2 of Evidence and Technological Authority

Mobile phone screenshots were used as evidence that the YouTuber had engaged with Momo. For example, in Figure *5-9* three screenshots of Momo (two of calls and one of

an exchange of messages with Momo) are presented side-by-side, documenting people's supposed interactions with Momo. Again, the succession of images creates a stronger sense of legitimacy – that there are multiple instances where people have interacted with Momo. The image of Momo is a salient feature, placed in the centre of the first two screenshots. The voice-over for these images negatively evaluates the Momo challenge as a '*massive disturbance*', tabling a 'Dangerous Momo bond'. There is also some TECHNOLOGICAL AUTHORITY, in the reference to '*social media*' and a '*messaging app*'. Again, the visual content reinforces the values legitimated in the voice over.



Authorisation: Impersonal: Technological ↘ Screenshots of mobile phone

Momo is a challenge that created a **massive disturbance** in social media recently. It features an avatar of a woman with dark hair, pale skin and oversized eyes, that asks people to add her to their contact list in a messaging app

Dangerous Momo bond ↘ [ideation: momo challenge/ attitude: negative REACTION]

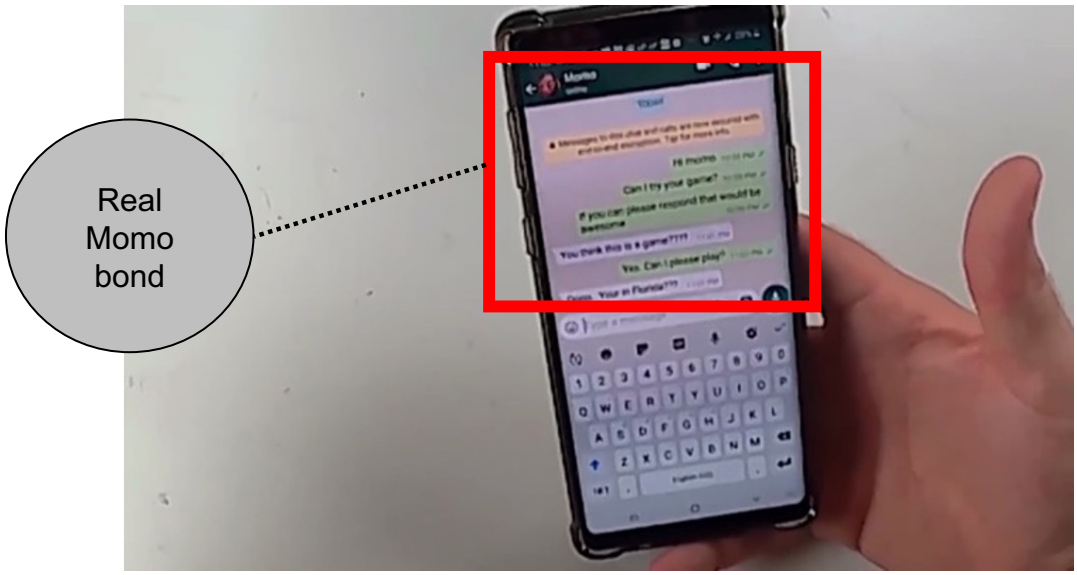
Temper: Modulate (massive disturbance)

Authorization: Impersonal: Technological ↘ '*social media*', '*messaging app*'

Figure 5-9 – Example 3 of Evidence and Technological Authority

The mobile phone as a physical item frequently featured in the Momo Challenge dataset, often with footage of people supposedly communicating with Momo via WhatsApp. The act of the YouTuber recording themselves in apparent real-time, with

the messages appearing on the mobile phone screen, construes TECHNOLOGICAL AUTHORITY, as if the ambient audience are witnessing an event as it occurs. However, anyone can change their WhatsApp name to 'Momo' and their profile picture to 'Momo' and set up this sort of staged interaction, as in Figure *5-10*. A 'Real Momo bond' is established via TECHNOLOGICAL AUTHORITY and the claim about '*the three messages I got*', again complements the visual legitimation for deceptive purposes.



Authorisation: Impersonal: Technological ↘ Mobile phone

…they **hung up on me** once the video chat didn't go through. That's what I'm guessing and I said: "you think this is a game", "oh you're in Florida ok", "I'll see you soon", that's the three messages I got from Momo.

Real Momo bond ↘[ideation: they (Momo)/ attitude: negative VALUATION]

Finesse: Embellish (I'm guessing)

Authorization: Impersonal: Technological ↘ '*three messages I got*'

Figure 5-10 – Example 4 of Evidence and Technological Authority

Overall, these examples show the importance of screenshots as evidence in the YouTube videos functioning in the service of legitimation. These screenshots were either construed as dynamic videos, collages of online videos, or filming in real-time the

YouTuber using their mobile phone. Despite the false nature of the Momo Challenge, evidence was still a central concept to these videos. Thus, these examples show how deceiving screenshots can be in these videos. From these examples, it can be argued that rather than purely believing the screenshot at face value, the social bonds that are attached to the screenshot need to be investigated. Social bonds that purely promote a real Momo or dangerous Momo bond without multiple layers of evidence (beyond a single screenshot) should be treated with caution.

5.3.5. **Moral Panics and Mythopoesis**

A final core theme revealed by the multimodal discourse analysis, was the spread of moral panics (Cohen, 1972) by the key presenters in the videos and how these moral panics were legitimated via acts of mythopoesis. As was previously highlighted, experts often spread moral panics about children using technology. However, this particular section will extend this analysis by discussing in further detail how presenters incorporated moral panics in their language and the key targets of their moral panic (for more explanation on moral panics as an academic theory refer to *The Literature Review* chapter).

In the videos, verbiage about the viral spread of the Momo Challenge, moral panics, or conspiracies were common. In the YouTube videos, discussions with parents and other family members were used frequently to support claims that the Momo Challenge was real. This would involve a parent directly stating that they had direct first-hand experience of the Momo Challenge:

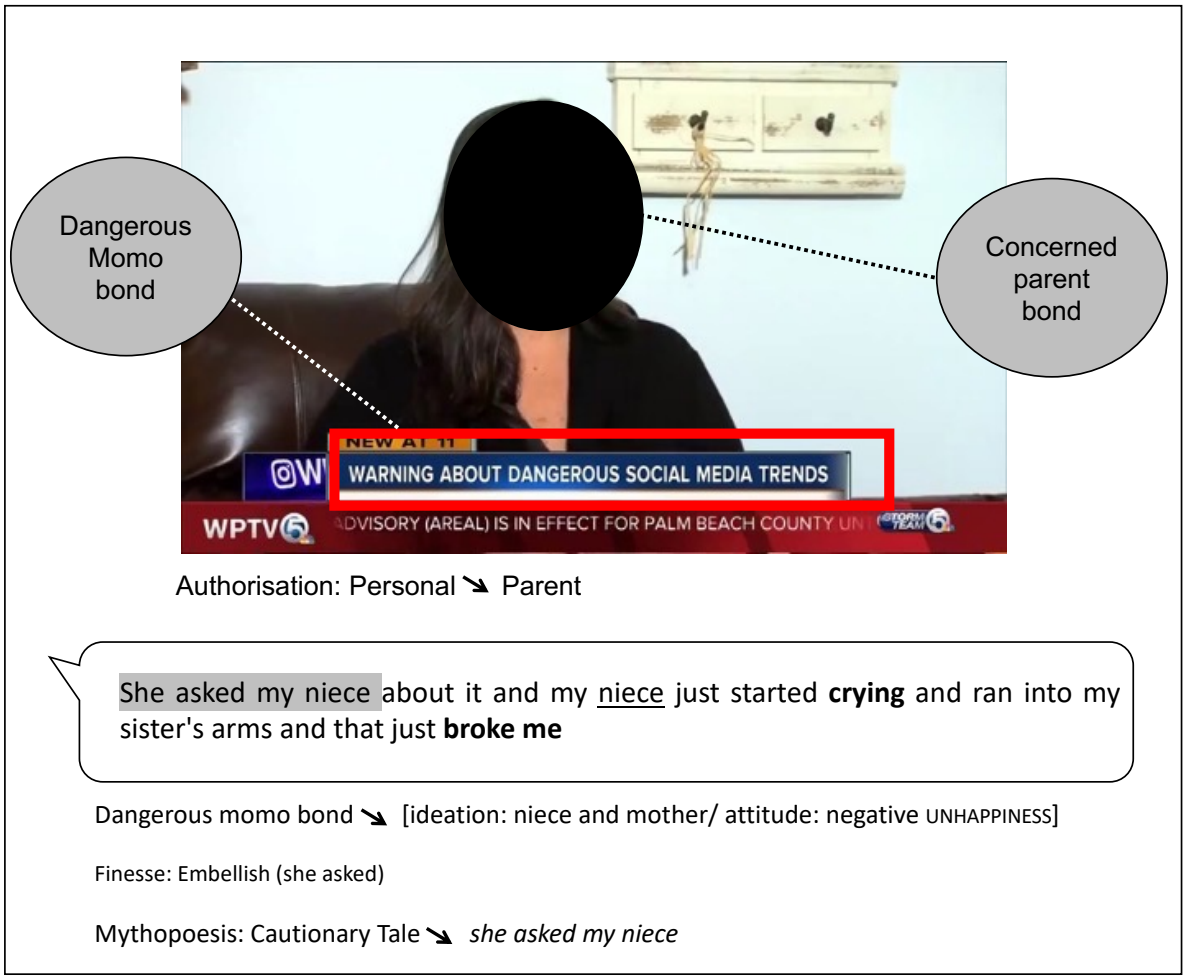
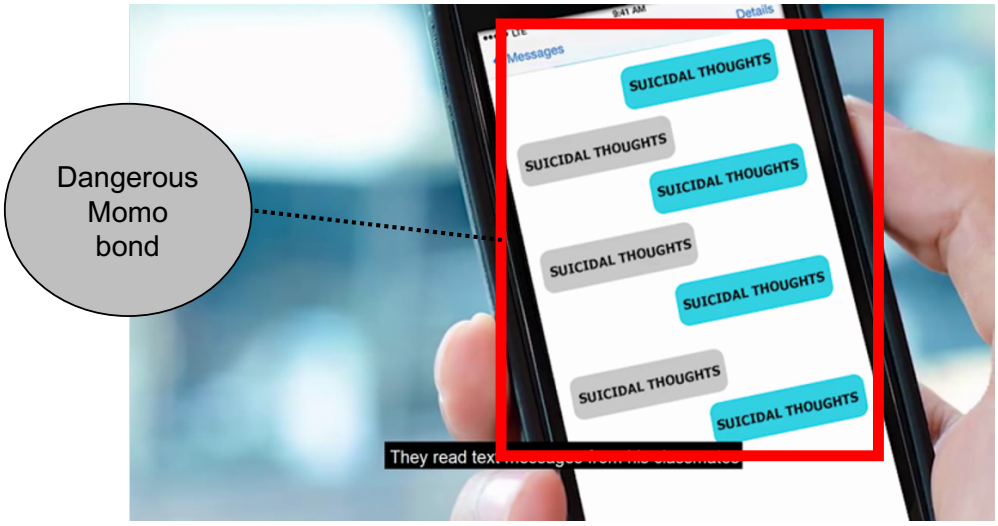


Figure 5-11 – Example 1 of Moral Panics and Mythopoesis

In this example, Momo is not an ideational target, instead the feelings of the niece are the focus as the niece is evaluated. Examples like this use recounts of past events to promote their claims about the realness and negativity of the Momo Challenge. These recounts were realised as MYTHOPOESIS, providing a cautionary tale about the dangers of the Momo Challenge with affective language ("*just started crying, broke me*"). Nonetheless, similar with the expert's legitimation strategy, there is a misalignment with the social bonds promoted. Whilst the family member in this example expresses a negative reaction towards the Momo Challenge, this negative reaction does not provide an inscribed evaluation of the Momo Challenge's veracity. Visually, the family member's face expresses a concerned parent/family member bond, coupled with another social bond 'the dangerous Momo bond' expressed via the dynamic banner running down the bottom of the screen, which features a negative evaluation of social media trends as

dangerous ("*warning about dangerous social media trends*"). The short nature of this segment in the news report additionally avoids further interrogation about the veracity of the Momo Challenge, rather the viewer just observes the emotional aspect of the story without being offered further validation. Thus, the power of this segment is in the moral panic it promotes through a personalised first-hand story with affective language.

In another example, screenshots were superimposed over a typical stock photo, albeit sometimes this method did not appear authentic as the editing was poor, as an attempt to illustrate a supposed recount of the effects of the Momo Challenge:



Authorisation: Impersonal: Technological ↘ Emulated screenshot of mobile phone

Before he died his parents noticed that the boy was frequently watching horror movies late at night. Sadly the **parents did not bother**, after **he died** that's when they found out that his son was doing the Momo challenge. They read text messages from his classmates **exchanging suicidal thoughts**.

Careless parents bond ↘ [ideation: parents/ attitude: negative CAPACITY]
Temper: Foster (frequently)
Evil Momo bond ↘ [ideation: Momo Challenge/ attitude: negative PROPRIETY]
Finesse: Distil (did not)
Authorisation: Impersonal: Technological ↘ *they read text messages*
Mythopoesis: Moral Tale ↘ *Before he died*

Figure 5-12 - Example 2 of Moral Panics and Mythopoesis

In this example the screenshot itself is not realistic but rather illustrative. The words '*suicidal thoughts'* are repeated multiple times in a font that is not associated with mobile phone texting interfaces. The voice-over negatively evaluates parents for not watching over their children, tabling a 'Careless Parents bond'. The Momo Challenge is also negatively evaluated for being associated with the child's death, contributing to an 'Evil Momo bond'. The claim by the voiceover that the parents '*read text messages from his classmates exchanging suicidal thoughts*' again presents the semiotic entity '*text messages*' as a legitimate source via the legitimation strategy of TECHNOLOGICAL AUTHORITY. Since this YouTuber does not actually have access to the original text messages, the emulated screenshot, with the repetition of the phrase '*suicidal thoughts',* is instead employed to bolster their claims.

All these examples share in common an emphasis of feelings over facts, and are short segments that do not allow the viewer to interrogate their claims of veracity. Judgemental language is directed towards a high-profile organisation or the media, but these judgements in particular were not attached to legitimation strategies. Overall, in these examples, a focus on conspiratorial content, the emotions of parents, and negative judgements about parenting or the internet, again moves discussion away from the facts of the Momo Challenge and into a more general discussion about internet dangers. Thus, the spread of moral panics and how easily they spread was represented via language features where the ideational target changes from Momo to broader societal issues and highly emotive language targeting other parents to adopt the shared social bond that the internet is dangerous.

5.3.6. **Summary of Results**

The results of this analysis will be focused on the key types of screenshots and screen recordings that emerged from this qualitative research. As screenshots and screen recording were frequently featured in the video dataset, they are the focus of this summary section. Table 5-1 illustrates the two most common legitimation strategies visually and linguistically for each video in the dataset. As this illustrates,

TECHNOLOGICAL AUTHORITY was the most common strategy for each video in the dataset either in the visual content of verbiage of the video, or across both.

Video	Most Common Legitimation Strategies (Visual)	Most Common Legitimation Strategies (Language)
1	Authorisation: Impersonal: Technological Authorisation: Commendation: Expert	Mythopoesis: Cautionary Tale Authorisation: Commendation
2	Authorisation: Impersonal: Technological Authorisation: Commendation: Expert	Authorisation: Commendation Mythopoesis: Cautionary Tale
3	Authorisation: Impersonal: Technological Authorisation: Personal	Authorisation: Impersonal: Technological Authorisation: Commendation
4	Authorisation: Impersonal: Technological Authorisation: Personal	Authorisation: Impersonal: Technological Moral Evaluation
5	Authorisation: Impersonal: Technological Authorisation: Personal	Authorisation: Impersonal: Technological Moral Evaluation
6	Authorisation: Impersonal: Technological Authorisation: Personal	Authorisation: Impersonal: Technological Moral Evaluation
7	Authorisation: Impersonal: Technological No other strategies	Authorisation: Impersonal: Technological Rationalisation: Explanation
8	Authorisation: Impersonal: Technological Authorisation: Personal	Authorisation: Impersonal: Technological Mythopoesis: Cautionary Tale
9	Authorisation: Impersonal: Technological Authorisation: Personal	Mythopoesis: Cautionary Tale Authorisation: Impersonal: Technological
10	Authorisation: Impersonal: Technological Authorisation: Personal	Authorisation: Impersonal: Technological Moral evaluation
11	Authorisation: Impersonal: Technological Authorisation: Personal	Authorisation: Impersonal: Technological Rationalisation: Explanation
12	Authorisation: Impersonal: Technological No other strategies	Authorisation: Impersonal: Technological (De)Authorisation: Impersonal: Technological
13	Authorisation: Impersonal: Technological Authorisation: Personal	Authorisation: Impersonal: Technological (De)Authorisation: Impersonal: Technological
14	Authorisation: Impersonal: Technological Authorisation: Personal	Authorisation: Impersonal: Technological Mythopoesis: Cautionary Tale
15	Authorisation: Impersonal: Technological Authorisation: Personal	Authorisation: Impersonal: Technological Mythopoesis: Cautionary Tale

Table 5-1 – Summary of Legitimation Strategies in the Dataset

Across the 15 videos, every example of a screenshot mentioned either verbally or visually was firstly identified. These were then categorised according to their verbal or visual attributes (as shown in the first column), and also taking into consideration the broader context of each screenshot and its interaction with other elements in the video. A summary of the types of screenshots, categorisations and frequencies across both case studies is shown in Table 5-2. As this table suggests, the most frequent screenshot visual structure was use of the split screen, followed by unaltered and emulated screenshots.

Screenshot visual structure	**Momo Challenge Dataset**	
	Frequency	**Type**
Evidence Collage	9 (1%)	Social media
Split Screen	254 (30%)	Social media, online articles split screen within split screen
Emulated Screenshot	127 (15%)	Social media, other technologies
Annotated Screenshot	22 (3%)	Social media
Unaltered Screenshot	130 (16%)	Social media, online articles
Frames that did not contain screenshots	291 (35%)	N/A
Total frames analysed	833	
Verbal reference to screenshots (number of times spoken about)	54	Social media, online articles, other technologies
Total transcript words	14 426	

Table 5-2 - Quantitative Results from Screenshots Analysis

Overall, three key strategies emerged from the analysis of the video transcripts; a focus on news stories and experts, using other YouTube videos and social media content as evidence for making claims about the Momo Challenge, and spreading moral panics about the Momo Challenge by providing warnings to parents and criticising YouTube.

While incorporating expert opinion and using evidence to back-up claims is considered good journalistic practice, the videos analysed often incorporated these practices deceptively e.g. interviewing an 'expert' who didn't actually understand the subject matter or manipulating content to make it appear as a trusted source. As was discussed throughout this section and annotated in the examples, a number of communing affiliation strategies were also employed in the dataset. Table 5-3 provides a summary of how these communing affiliation strategies were distributed throughout each video, with the most common strategy highlighted in each row in grey. Video 6 is an outlier in this table because it had minimal communing affiliation due to its structure being an interview, hence dialogic affiliation strategies were more common. With the rest of the dataset, the majority of videos heavily featured MODULATING affiliation strategies as social bonds were emphasised in order to make them more engaging. In similarity to the typical vlog structure, MARSHALLING affiliation strategies were also common as YouTubers typically addressed their imagined audience in their monologues. Lastly, EMBELLISHING was usually a more common strategy than DISTILLING in the dataset, due to the speculative nature of the Momo Challenge, where YouTubers entertained numerous theories about the Momo Challenge. This summary briefly maps out the range of communing affiliation strategies used to position the social bonds explored in this chapter.

Video	Total Couplings Identified	MARSHAL (directly addressing a person or community)	DESIGNATE (naming a community)	MODULATE (adjusting scope of venture)	FOSTER (adjusting degree of venture)	DISTIL (making a definitive statement)	EMBELLISH (entertaining other possibilities)
1	19	11%	16%	32%	0%	5%	36%
2	18	0%	6%	39%	6%	6%	43%
3	14	7%	14%	43%	0%	0%	36%
4	54	11%	8%	47%	0%	4%	30%
5	21	28%	10%	29%	5%	14%	14%
6	130	Interview structure – dialogic affiliation strategies most common					
7	29	10%	0%	38%	4%	24%	24%
8	21	5%	5%	42%	5%	10%	33%
9	10	28%	0%	38%	14%	0%	20%
10	29	38%	6%	33%	3%	3%	17%
11	46	15%	2%	57%	2%	11%	13%
12	30	13%	3%	43%	0%	26%	15%
13	32	9%	9%	31%	0%	18%	33%
14	29	17%	3%	59%	0%	10%	11%
15	31	16%	3%	35%	0%	16%	30%

Table 5-3 – Communing Affiliation Strategies across the Transcript Dataset

5.4. Illustrative Examples from the Dataset

This section will provide detailed illustrative examples of two distinct macro-genres identified in the dataset: the news reporting macro-genre and the entertainment macro-genre. It will discuss the key features of these macro-genres, and how different legitimation strategies are used throughout the videos in order to legitimatise the key social bond that the Momo challenge is a real and dangerous threat. Specific examples have been selected in order to show how the affiliation and social bonding analysis works across the logogenetic unfolding of videos representing these particular macro-genres. This logogenetic perspective offers a contrast to the synoptic perspective of the previous section which focused on trends in affiliation and legitimation strategies across the entire dataset.

5.4.1. **News Reporting Macro-Genre**

The news-reporting macro-genre videos incorporated similar legitimation and genre stages to the other macro-genre videos to highlight the importance of news stories and expert interviews regarding the Momo Challenge. This structure involved an initial introduction of the Momo Challenge by providing an attitudinally rich description of Momo, as in the following examples:

1. Yes it is a very **strange** name but it is very **disturbing** and it's important to talk about so this challenge involves people taking commands from unknown numbers and social media accounts for 50 days and the part that's extremely **concerning** is the **dangerous** tasks that the so called Momo character is asking players to perform.
   [ideation: momo/ attitude: negative PROPRIETY]
   Temper: Modulate (very)
   Dangerous Momo bond
   [ideation: momo character/ attitude: negative PROPRIETY]
   Temper: Modulate (extremely)
   Dangerous Momo bond

2. Originally a Japanese statue, this is Momo with **bulging** eyes, a **chilling** smile and jet-black hair on a bird's body.
   [ideation: momo/ attitude: negative REACTION]
   Finesse: Distil (this is)
   Dangerous Momo bond

3. She's got bug eyes, long stringy hair, a **monstrous grimace** and **spooky** chicken legs. Imagine this **freaky** character suddenly popping up in cartoons on YouTube channels for kids.
   [ideation: momo/ attitude: negative REACTION]
   [ideation: momo appearing on YouTube/ attitude: negative PROPRIETY]
   Convoke: Marshal (imagine this)
   Dangerous Momo bond

All these examples focus on Momo as the main ideational target, evaluating 'Momo' (as a person) with negative PROPRIETY, or evaluating the 'Momo Challenge' (as an object)

with negative REACTION. These attitudes establish a 'dangerous Momo bond'. All the news story genres begin with a dramatic account of the Momo Challenge, aiming to get audiences worried about Momo. TEMPERING affiliation was a common strategy to draw attention to these details about Momo's appearance, or CONVOKING affiliation, marshalling people to 'imagine' what Momo is.

The videos in the news-reporting macro-genre can also be analysed visually as multimodal documents, with multiple layers of dynamic meaning. By considering Bateman's concepts of 'page-flow', 'text-flow' and 'image-flow', the way in which multiple elements work together in a single video frame can be explored (Bateman, 2008; Bateman, 2013). For example, in videos that belonged to the news-reporting macro-genre (such as *Figure* 5-13) a common example included:

- Image Flow – Background images were often moving, or foreground images were animated.

- Text Flow – A banner at the bottom of the screen, with text such as "Warning to Parents About the 'Momo Challenge' ", adding an extra element of negative PROPRIETY.

- Page Flow – Moving from frame to frame, the visual content would often switch from different social media posts, with the newsreader remaining in the foreground.

Figure 5-13 – Example of Flow in the News Reporting Macro-genre

This shows that screenshots are an important part of the flow process, contributing to an information overload, where the viewer does not focus on a single piece of information, rather navigating multiple pieces of information and different intensities of flow. Thus, this abundance of information is effective in instances of information disorder, by creating content that appears legitimate and that the viewer does not have time to think about its actual veracity.

After providing some context about the Momo Challenge, interviews were conducted with an "expert" – ranging from a principal at a high school, a psychologist, or a technology or cyber-security expert. The ways in which experts were introduced included:

4. Narrator: So who and what is Momo? Cyber security **expert** (name):
   [ideation: cyber security expert/ attitude: positive CAPACITY]
   Finesse: Embellish (Cyber security expert)
   Knowledgeable expert bond

5. Newsreader: Technology **experts** say this is the opportunity for parents to remind their kids **not to trust unknown numbers** and to not click on unidentified links. Experts also encourage you to **change** email and social media passwords very frequently and **block** unknown numbers inviting you to play right away.
   [ideation: technology expert/ attitude: positive CAPACITY]
   Knowledgeable expert bond

[ideation: kids/ attitude: negative CAPACITY]
Ignorant kids bond
Finesse: Embellish (Technology experts say)
[ideation: unknown numbers & links/ attitude: negative REACTION]
Dangerous unknown numbers bond
Finesse: Embellish (Experts also encourage)
Convoke: Marshal (you)
Temper: Modulate (very)

Despite these experts being introduced as knowledgeable in the field, often their knowledge did not translate to the actual concerns of the Momo Challenge. For example, technology experts were more focused on reminding people to change their passwords rather than the details of the Momo Challenge. Momo is not the ideational target, instead it is the notion that parents should learn more about technology and teach their kids about technology.

To summarise the linear structure of the news-reporting macro-genre, a typical genre staging structure consisted of news anchors announcing the story, and a reporter then describing the Momo Challenge using a lot of evaluative language and social media as evidence. Interviews with either an 'expert' or concerned parent/teacher were then conducted, before showing 'evidence' of a school's letter warning people about the Momo Challenge or a warning shared by a celebrity. Lastly, the Momo Challenge was linked to a 'supposed' suicide, with again social media posts being used as evidence, and a warning to parents to watch over their children. This structure is shown as a diagram in Figure 5-14. As this diagram shows, there are three key stages of the news-reporting macro-genre video: the newsroom as the introductory stage, the news story with the reporter voiceover as the significant part of the video, followed by the video returning to the newsroom for the concluding remarks. Within these three key stages, there are multiple genre stages that introduce particular elements or participants in the story and provide evidence to back-up propagated social bonds (that Momo is a real and tangible threat).

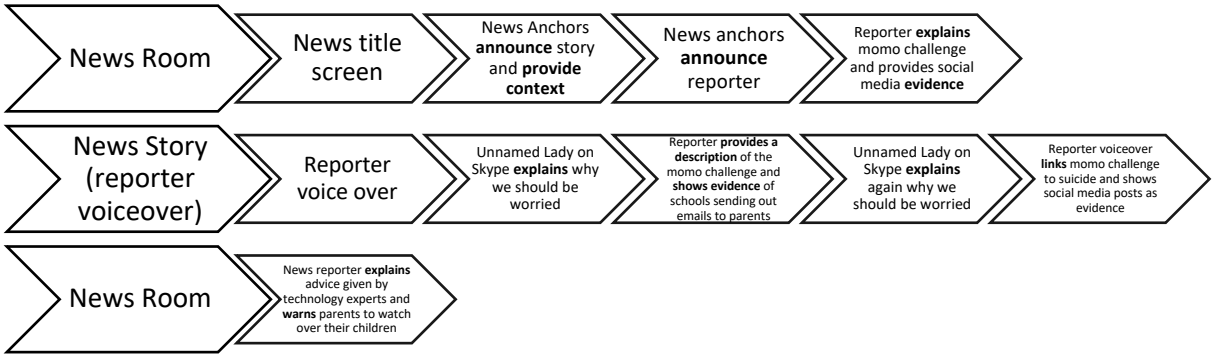
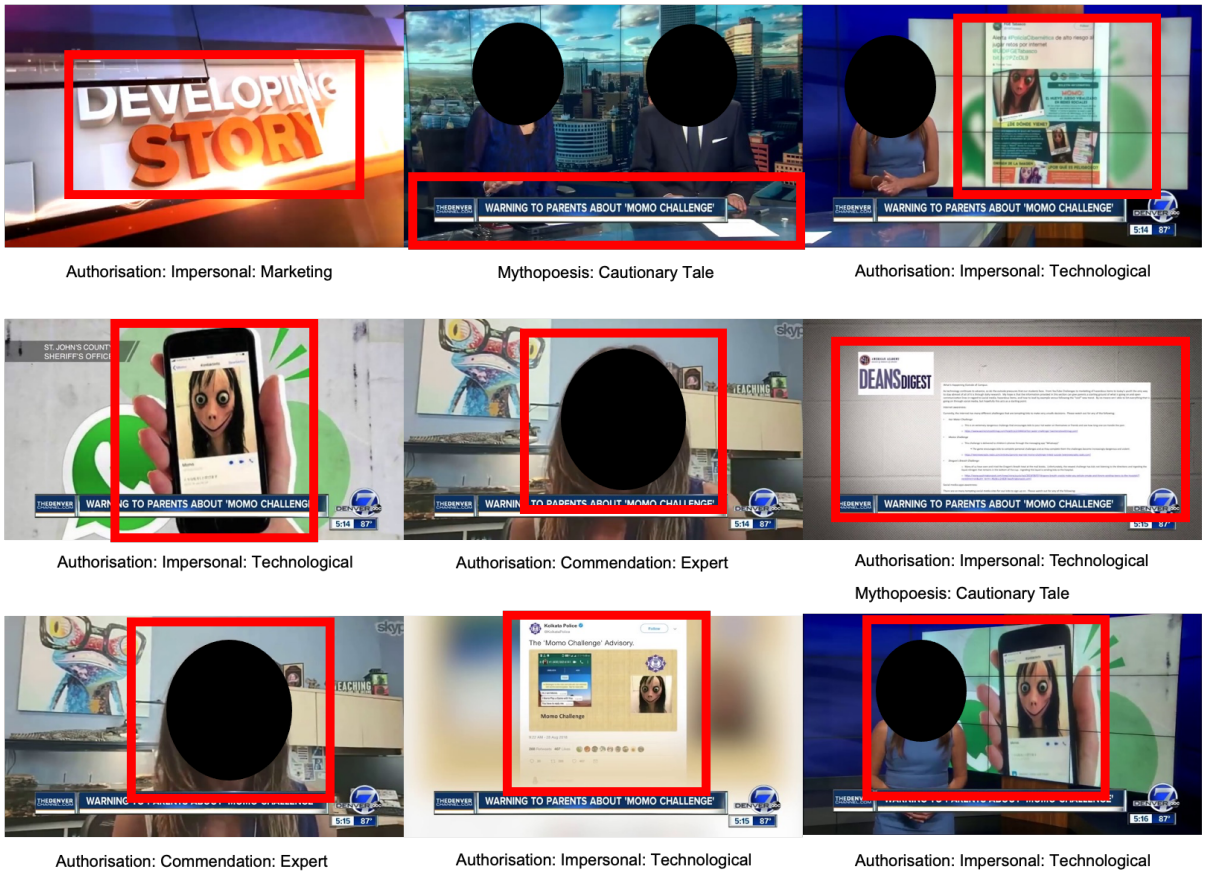


Figure 5-14 – Genre Structure for a News Reporting Macro-Genre Video

Visually, there were several key legitimation stages throughout the news-reporting macro-genre video that can be summarised. Firstly, MARKETING AUTHORISATION was evident in the introductory footage of the video, either promoting a 'developing story' or featuring the news station's logo. When footage turned to the news anchors and reporters, the dynamic banner at the bottom of the screen, would represent MYTHOPOESIS: CAUTIONARY TALE by the negative attitudinal language providing a warning about the Momo Challenge. TECHNOLOGICAL AUTHORITY was the most common legitimation strategy throughout the video, as seen by multiple visual references of screenshots shown on the news station projector screen, or as a full shot image accompanied by the reporter's voice-over. EXPERT AUTHORISATION was the second most common legitimation strategy, with short segments devoted to an interview with an authority figure. These strategies are summarised in Figure 5-15.



Authorisation: Impersonal: Marketing | Mythopoesis: Cautionary Tale | Authorisation: Impersonal: Technological

Authorisation: Impersonal: Technological | Authorisation: Commendation: Expert | Authorisation: Impersonal: Technological

Mythopoesis: Cautionary Tale

Authorisation: Commendation: Expert | Authorisation: Impersonal: Technological | Authorisation: Impersonal: Technological

Figure 5-15 – Key Visual Legitimation Strategies in the News Reporting Macro Genre of Videos

Overall, the news reporting macro-genre was characterised by a formulaic approach to constructing interest and credibility regarding the Momo Challenge. The main social bond the news reporting macro-genre tried to construct, as a 'dangerous Momo bond'. This bond was primarily legitimated via visual strategies in TECHNOLOGICAL AUTHORISATION and EXPERT AUTHORISATION, accompanied at times by references in language that evoked MYTHOPOESIS. These strategies will now be compared to another distinct macro-genre in the dataset: the entertainment macro-genre.

5.4.2. Entertainment Macro-Genre

The entertainment macro-genre focused on entertaining audiences about the Momo Challenge by either trying to call Momo or reacting to videos about Momo, rather than following a typical news report structure. With all the videos in the dataset that didn't

belong to the news-reporting macro-genre, the treatment of 'experts' and 'new stories' was structured differently. These videos relied less on using evidence from experts, but highlighted the importance of the Momo Challenge by linking it to recent news coverage:

6. Hey what's going on everybody welcome back to the channel! If you're new here my name is (name), in today's video I want to talk about <u>something</u> that is a very very **hot** topic on YouTube right now and <u>it</u> should be….
   [ideation: Momo challenge/ attitude: positive VALUATION]
   Important Momo bond
   Convoke: Marshal (hey what's going on everybody)
   Temper: Modulate (very very)

7. What's up guys? How's it going? So anyways I'm pretty sure you guys have scrolled through facebook, Instagram and saw this Momo challenge thing. It's not, <u>it's</u> **nothing new**. It's been going on for like since last year, but this year <u>it's</u> **starting to surface upon the news** and others are more aware of these challenges because now it's just slipping through the cartoon.
   [ideation: Momo challenge last year/ attitude: negative VALUATION]
   Dangerous Momo bond
   [ideation: Momo challenge this year/ attitude: positive VALUATION]
   Dangerous Momo bond
   Convoke: Marshal (what's up guys, you guys)
   Finesse: Distil (nothing new)
   Temper: Modulate (more aware, just)

In these videos, the Momo Challenge was introduced by referring to its presence in mainstream news and its current virality. CONVOKING affiliation was initially used in order to address those watching the video. In these cases, the Momo Challenge was evaluated with POSITIVE valuation – as something worthwhile to be making a video about. Thus, compared to the news reporting macro-genre, others macro-genres had a more conversational tone, addressing the audience directly.

In some cases, the entertainment macro-genre also referred to experts in order to give their videos credibility. For example, in this case, experts such as psychologists are treated with positive CAPACITY:

8. Robert B Calvini, a **respected** social scientist and specialist in the area of compliance psychology says the automatic stereotyped behavior is prevalent in much of human action, he cites an experiment by Harvard social psychologist Ellen Langer where you can see the concept in action.
    [ideation: Robert B Calvini/ attitude: positive CAPACITY]
    Knoweldgable expert bond
    Finesse: Embellish (says)

Similar to the news reporting macro-genre, an expert's opinion is used, despite that opinion not using the Momo Challenge as a main ideational target. In this example, Calvini's theory of compliance psychology is linked to the Momo Challenge by the YouTuber, as an attempt to give the video some credibility. This is a distraction from questions regarding the Momo Challenge's veracity.

Another way of giving a video credibility, involved using schools as trusted sources. For example, in this text, Momo is treated with negative PROPRIETY based on schools issuing warnings:

9. Today schools across Britain began issuing **warnings** on their websites and social media accounts saying they have been contacted by hundreds of **concerned** parents. Challenges appear midway through kids YouTube, fortnight and peppa pig to **avoid detection** by adults.
    [ideation: challenges (momo)/ attitude: negative PROPRIETY]
    Dangerous Momo Bond
    Temper: Foster (hundreds of)
    Finesse: Embellish (saying)
    [ideation: adults/ attitude: negative CAPACITY]
    Ignorant parents bond
    Finesse: Distil (to avoid)

The main affiliation strategy here is a TEMPERING affiliation strategy 'hundreds of concerned parents' and a FINESSING affiliation strategy 'saying' expanding the coupling and 'to avoid' contracting the coupling. In the second sentence, similar to the interviews with the experts, parents are treated with negative CAPACITY, due to their inability to understand technology. Thus, with warnings from a trusted institution (schools) and the references to concerned parents, this legitimises the significance of the Momo Challenge.

Visually, the entertainment macro-genre had varying levels of 'flow', but each still represented the same notion of multi-layered documents pretending to act as legitimate sources. In the entertainment macro-genre videos, the 'flow' was less predictable. One example is the image flow consisting of the presenter and a YouTube clip, an absence of text flow, and the page flow consisting of the YouTube clip changing more rapidly than the video of the presenter reacting. Examples of this are shown in *Figure 5-16* and *Figure 5-17*. Another example is in videos that were more animated (abstract and computerised coding orientation) rather than naturalistic. The 'flow' in these videos often consisted of blurred images changed in the background, text in the centre of the frame and the text changing or fading out as part of the page flow, as shown in *Figure 5-18*.

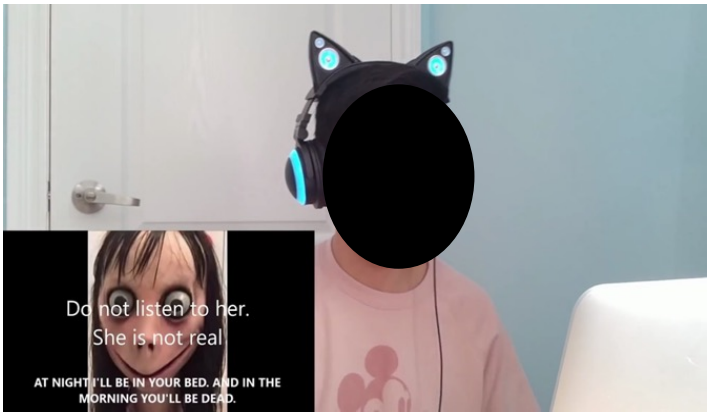


Figure 5-16 – Example 1 from Entertainment Macro-Genre

Figure 5-17 – Example 2 from Entertainment Macro-Genre

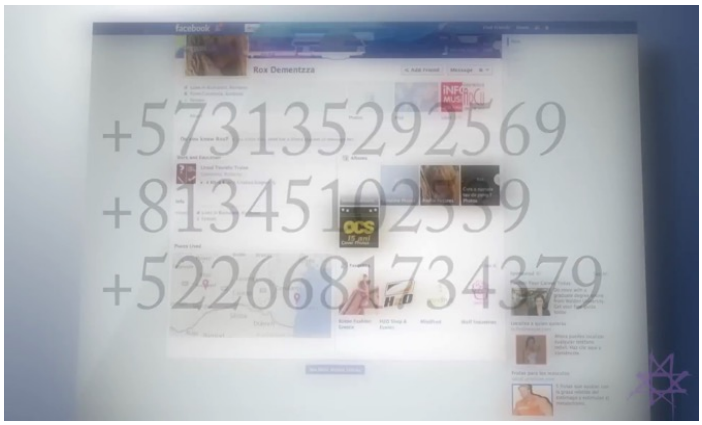


Figure 5-18 – Example 3 from Entertainment Macro-Genre

In the videos, discussions about the viral spread of the Momo Challenge, moral panics, or conspiracies were common. In terms of evaluative meaning, YouTubers would focus their evaluations on the target of the conspiracy or moral panic. There was also discussion in the videos about why Momo went viral, for example:

10. So, while it's a little unclear when this started happening, <u>this</u> has **blown up online** with many people making videos around messaging the character and trying to explain the phenomenon. As an online creator yourself, why do you think things like this spread online?
    [ideation: this (momo challenge)/ attitude: positive VALUATION]
    Trending Momo bond

    i. It's the same way conspiracy theories blew up online. Everyone **likes** <u>things that are unknown</u>.

[ideation: unknown things/ attitude: positive VALUATION]
Positive conspiracy bond
Manage: Support: Defer

In this dialogue, between the presenter and interviewee, the interviewee likens the Momo Challenge to a conspiracy theory in terms of its virality, and positively evaluates 'unknown things'. This exchange between the presenter and interviewee, is interesting for the reason that the unknown nature and viralness of the Momo Challenge is highlighted.

In the YouTube videos, parents were used frequently to support claims that the Momo Challenge was real. This would involve a parent directly stating that they had first-hand experience of the Momo Challenge:

11. That thing telling her to throw herself in the garbage and to throw herself from, um, from a slide, to hurt her head that's just **terrifying**.
    [ideation: that thing (Momo)/ attitude: negative REACTION]
    Dangerous Momo bond
    Promote: Modulate (just)

12. My kids are watching just their shows and little video clips are popping up and doing this stuff and you have to guard your kids with your own life because if you don't then **something bad can happen to your own kids**. So if you're not aware of this go check it out I hope I brought attention to some of y'all that **may not know about this** and your kids haven't came across it yet, maybe you can prevent them from doing that and um but that being said do take the proper steps to fix this.
    [ideation: videos/ attitude: negative VALUATION]
    Dangerous videos bond
    [ideation: parents who do not know about momo/ attitude: negative CAPACITY]
    Ignorant parents bond
    Convoke: Marshal (you have to, if you're not aware & some of y'all)
    Finesse: Embellish (maybe)

In the first example, TEMPERING affiliation is used, with the use of 'just' emphasising a particular point. In the second example, YouTube videos are treated with negative VALUATION and parents that don't know about Momo have negative CAPACITY. In this example, CONVOKING affiliation is predominately used, as the parent is directly addressing other parents, warning them about the Momo Challenge. All these examples use anecdotes to promote their claims about the realness and negativity of the Momo Challenge.

Affiliation strategies were also useful in identifying how intertextuality and evidence were embodied as themes. FINESSING affiliation was useful in identifying how the presenter treated the intertextual reference as something they embodied or something they were distanced from. CONVOKING affiliation signalled that the presenter wanted their audience to join in with them, in terms of finding evidence to back-up the Momo Challenge. TEMPERING affiliation would strengthen the claims of the YouTuber and draw attention to evaluated entities. Overall, a mix of affiliation strategies were used, with each expressing different motives that the presenter may have had in getting users to bond with their claims.

A typical genre-structure for macro-genres that belonged to either the entertainment, advertising, educational, or commentary genres involved an introduction composed of a title screen and company logo. 'Found' footage or clips from 'news' footage of Momo would be shown, as evidence of the Momo Challenge existing. A presenter would introduce the topic and provide context about the Momo Challenge. Discussion about the Momo Challenge would then take place and/or the presenter would try to show how the Momo Challenge works. The last genre category would involve the presenter asking the audience for their thoughts on the challenge and encouraging them to subscribe to their channel. A visualisation of this type of macro-genre structure is shown in Figure 5-19.

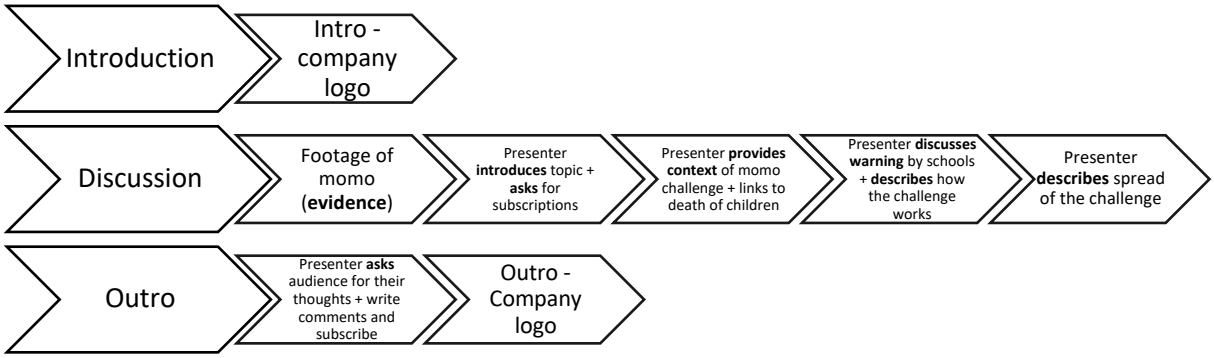


Figure 5-19 - Genre Structure for the Entertainment, Advertising, Educational, or Commentary Macro-Genres

In terms of visual legitimation strategies for the entertainment macro-genre, there was a particular structure to how these legitimation strategies featured throughout the video. At the start of the video, MARKETING AUTHORISATION was used to promote the video channel. A video collage of different YouTube videos was then featured, primarily containing footage from news stations (videos that were also included in this dataset), this can be realised as an example of TECHNOLOGICAL AUTHORITY. This TECHNOLOGICAL AUTHORITY also existed once the video turned to the YouTuber's vlog, with the physical mobile phone that the YouTuber frequently holds up to the screen acting as their supposed evidence for the Momo Challenge occurring. This was accompanied with further supposed evidence of the Momo Challenge spliced videos appearing in YouTube search results. These examples can be seen visually in Figure 5-20.

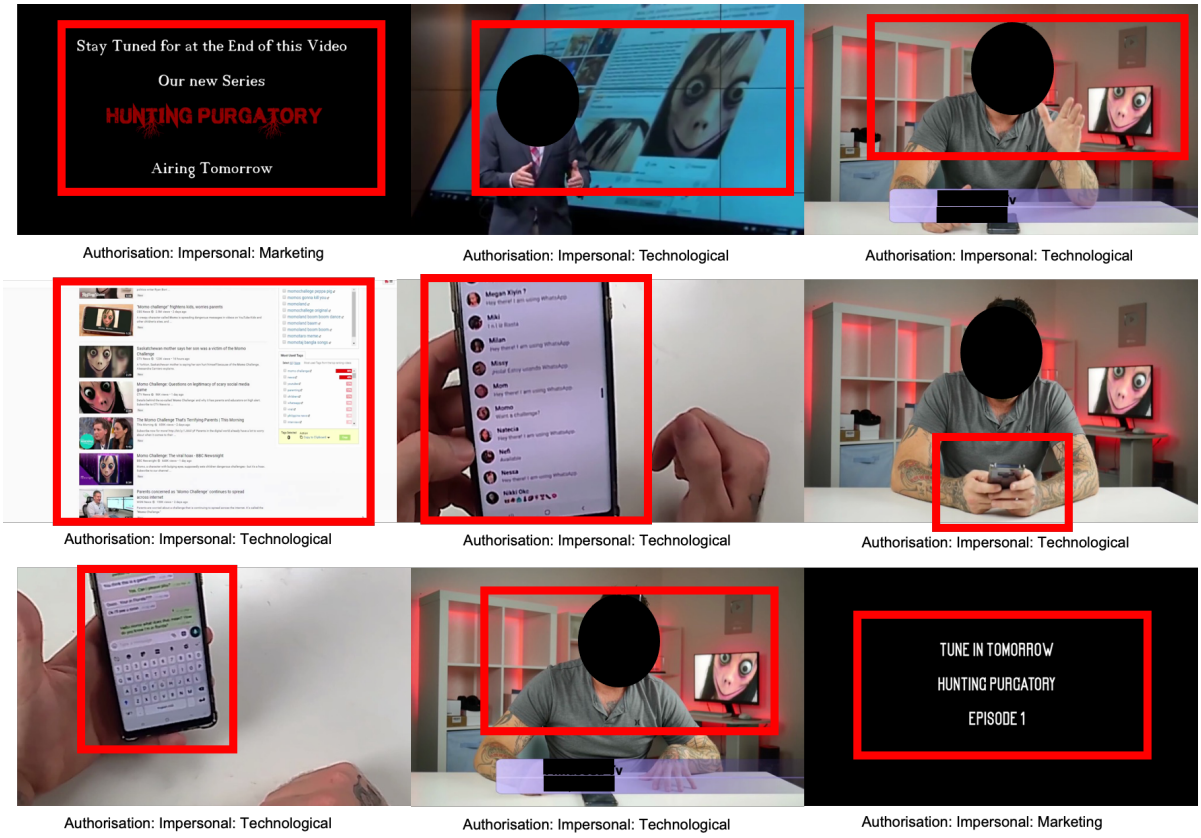


Figure 5-20 - Key Visual Legitimation Strategies in the Entertainment Macro Genre of Videos

Similar to the news reporting macro-genre, TECHNOLOGICAL AUTHORITY was the main legitimation strategy to propagate the bond that Momo is a real and tangible threat. Nonetheless, whilst the news reporting macro-genre also relied to experts, the entertainment macro-genre instead turned to the PERSONAL AUTHORISATION of the YouTuber in the vlog (regardless of their lack of expertise as a figure). Rather, their expertise existed in their mobile phone device and first-hand account of contacting Momo. As this examples shows, multiple layers of legitimacy are needed to spread the social bond that Momo is a real and tangible threat.

12.1. Conclusion

YouTube videos can spread conspiratorial and hateful content in sophisticated ways, and screenshots contribute to the believability and virality of misinformation and disinformation. This chapter has demonstrated how a social semiotic approach can

illuminate the affiliative and legitimation function of verbal and visual evidence in deceptive YouTube videos. My addition of TECHNOLOGICAL AUTHORITY to the legitimation framework was made in order to account for the importance of technology as evidence in social media discourse. This study also connects to the broader research currently being conducted on screenshots that encourages researchers to look at the screenshot itself as a media object or a kind of document, rather than just looking at the content of the screenshot and ignoring the now societal importance of screenshots as artefacts in our everyday lives (Frosh, 2018).

The visual modality and legitimation analysis undertaken in this chapter shows how texts are made to look authentic, and therefore believable. As the examples showed, this visual content was deceptive as it did not align with actual online encounters. In addition, whilst the transcript texts were also deceptive, the visual content often added to the meanings of the transcript, rather than simply illustrating the verbiage. Thus, the continual imagery of screenshots and mobile phones was an important aspect in the effectiveness of the videos and their believability. Overall, an understanding of 'flow' in the video dataset shows how the videos act as multi-modal documents that are persuading the audience that their content is real; multiple visual, textual and verbal content happening all at once that can be compared to the concept by Krafft and Donovan (2020) of 'evidence collages'. All the videos have been thoroughly edited, with as much visual content crammed in as possible. This abundance of elements in a frame all flowing at once, shows how a profusion of the visual can contribute again to the deceptive viral appeal of these videos, by pretending that this content is legitimate. When multiple screenshots are shown on a screen, often the viewer cannot think about the elements singularly, but rather believes the content at face-value because they are being bombarded by screenshots and video clips – again, often thought of as legitimate content. Therefore, the abundance of information and false legitimacy are interconnected.

This chapter has also explored the role of deceptive genres of communication and moral panics in YouTube videos about The Momo Challenge. It has analysed how genre, affiliation and legitimation unfolds in the transcripts to construct deceptive narratives

that propagate social bonds relating to moral panics associated with children using technology and the dangers of the YouTube platform. The analysis of the visual content in the videos, further supported the claims that technology is being used to elicit moral panic. In terms of social bonds, across the videos there was a focus on broader criticisms of parenting skills, children using technology, and the YouTube platform.

To conclude, this case study has shown how an affiliation and legitimation analysis can be used as strategies for understanding what 'evidence' YouTuber's use to spread their opinions and form social bonds to rally their audiences around their work. In relation to the overarching research question, a systemic functional linguistics approach provides a framework for understanding how the bonds that users promote in their discourse, determines their relationship to deceptive content and their engagement with moral panics. In the case of the Momo Challenge, users are bonding around more than just whether Momo is real or fake. These users are bonding around additional clusters regarding concern about parenting in terms of managing digital material (bad YouTube bond), worries over generational shifts (bad older people bond), and concern with establishing cultural dominance (superior nation/knowledge bonds). Thus, employing the Momo challenge as a case study illustrates how an affiliation and legitimation analysis can be used to understand *why* people believe false information on social media through their alignment with particular social bonds (the interpersonal realm), not just identifying *what* is false (the ideational realm).

6. CHAPTER SIX THE NOTRE DAME FIRE: POLITICALLY MOTVIATED CONSPIRACIES AND WHITE SUPREMACIST DISCOURSE

6.1. Introduction

This chapter explores a case study that involves political misinformation and conspiracy theories, in comparison to the previous chapter on the Momo Challenge case study which focused on misinformation arising from moral panics about technology. The overarching focus of this chapter is on how social bonds in discourse are legitimated or delegitimated, in language and visually. This chapter will firstly provide some background information on the Notre Dame Fire, and the nature of political misinformation and conspiracy theories. It will then turn to the affiliation and legitimation analysis that was conducted on the video transcripts and visual content of the videos. Lastly, the main results from this case study will be reflected upon and future directions discussed.

6.1.1 The Notre Dame Fire, Political Misinformation and Conspiracy Theories

The Notre Dame Cathedral fire occurred on 15th April 2019 in Paris, France. This fire caused the Cathedral's spire to collapse and extensive damage to the roof and upper walls of the building. French prosecutors declared that there was no evidence of the fire being a deliberate act, with the fire most likely being caused by an electrical fault. Despite the fire occurring in Paris, this news event received a global reaction and was one of the most googled news events of 2019 worldwide[20]. Due to this global impact, conspiracy theories about how the fire started proliferated on social media, for example, in posts blaming Muslims for starting the fire, and labelling it as a terrorist act. Just

[20] See: https://trends.google.com/trends/yis/2019/GLOBAL/

shortly after the Notre Dame Fire had started, news organisations such as Buzzfeed, The Guardian and CNN were reporting the internet being 'awash in Notre Dame conspiracies'[21]. It is this global attention that the Notre Dame Fire attracted that makes it worthy to consider as a case study for this thesis.

Social media platforms such as YouTube have been associated with far-right extremists, white supremacists, and conspiracy theorists. When the Notre Dame Fire occurred, this viral news story was seized by these far-right extremists, white supremacists, and conspiracy theorists. In particular, YouTube was a platform where people could post videos while an investigation into the fire was still ongoing, spreading misinformation about who started the fire, deceptively using the global event to direct people's anger towards migrants in France, or eroding trust in mainstream media organisation's coverage of the fire. Social media platforms such as YouTube are also amplifiers and manufacturers of racist discourse due to the affordances, business models and cultures that they maintain through forms of 'platformed racism' (Matamoros-Fernández, 2017). In a similar way, YouTube is platforming conspiracies. Research has shown that conspiracy theories are linked to racist and xenophobic discourses, for example, in the form of islamophobia (Shooman, 2016; Farkas et al., 2018), or antisemitism (Allington et al., 2021) (refer to *The Literature Review Chapter* for further discussion for these concepts).

This chapter has three core aims:

1. To examine the transcripts of videos that spread misinformation about the Notre Dame Fire, in order to understand the main values that the YouTuber adopts in order to affiliate with their audience, and the (de)legitimation strategies they use in order to provide evidence for their claims

[21] See: https://www.buzzfeednews.com/article/janelytvynenko/notre-dame-hoax-timeline; https://www.theguardian.com/world/2019/apr/16/social-media-platforms-failed-to-counter-notre-dame-fire-conspiracies-say-critics; https://edition.cnn.com/2019/04/16/tech/conspiracy-theories-notre-dame-cathedral-fire/index.html

2. To examine the visual content of the videos from a multimodal discourse perspective, in order to understand if visual evidence provides greater legitimation of a YouTuber's claims and to understand the affiliation potential of visual content
3. To develop a method that shows the congruence of language and visual content in the formation of social bonds and the legitimation of these social bonds

Building on the methods used in the Momo Challenge case study, this chapter presents an adapted legitimation framework (Van Leeuwen, 2008; Van Leeuwen, 2007) that complements the appraisal (Martin and White, 2005) and affiliation analysis (Zappavigna and Martin, 2018; Zappavigna, 2018). In other words, this chapter explores how social bonds in discourse are legitimated, in language (by analysing video transcripts and comments) and visually (by analysing the visual content of the videos). This method is useful for understanding the linguistic and multimodal strategies used in conspiratorial discourse and to apprehend that conspiratorial discourse cannot be linked to one specific profile, rather, there are various linguistic identities that create and engage with conspiratorial discourse, with their own unique array of legitimation and affiliation strategies.

6.2. Content Analysis of Video Metadata

This section will provide an overview of the video dataset and the macro-genres unique to each video. For a detailed account of the data sampling strategy and criteria for selection please refer to *The Methodology Chapter*. To briefly summarise, the 15 videos selected were in English, had the highest number of views compared to other conspiratorial videos, had comments enabled, and were created in the 24-hour period after the Notre Dame Fire occurred on 15th April 2019.

6.2.1. Video Macro-Genres

The videos represented a range of different conspiracy theories, ranging from misinformation (unintentionally sharing false information, i.e. believing that one is telling the truth) to disinformation (intentionally sharing false information, often with a political motive) (Wardle and Derakhshan, 2017). Ideationally, the videos all construed experience via enacted participants and circumstances. Represented participants included presenters, screenshots and centred text. Videos were dynamic and involved speech. Circumstances ranged from virtual backgrounds to the presenter located outside the Notre Dame Cathedral, or in a room. The targets of the conspiracies included Muslims (either blaming Muslims for causing the fire or directing anger towards Muslims apparently expressing happiness towards the Notre Dame Fire), President Emanuel Macron, French migrants, authorities (encompassing French Prosecutors or French Firefighters), Freemasons, Globalists, or well-known conspiracies linked to UFOs or the New World Order. From this list of examples, we can see there was a diverse array of targets, with videos expressing recycled conspiracy theories, distrust towards authorities, or contemporary xenophobic or racist discourse.

In terms of video structure, the videos had several different formats. For example, nine of the videos were vlogs (video blogging, a video structure where the YouTuber is directly speaking towards the camera). Vlogging is very common among YouTube videos (Burgess and Green, 2018), so this strategy was not surprising. Whilst three of the videos only incorporated the YouTuber directly speaking to the camera with minimal multimedia content, the other six vlogs heavily featured multimedia content, either switching back and forth between the vlog and social media content, or simultaneously showing the vlogger in one corner of the screen and social media content or a live-stream taking up the rest of the screen, or vice-versa. Bateman's concept of 'page-flow', 'text-flow' and 'image-flow' explains the ways in which these multiple elements work together in a single frame (Bateman, 2008). For example, a common structure for a YouTube video involved in the corner of the video frame a vlog of the YouTuber, and the rest of the video frame consisted of dynamic computer screen featuring a live stream or news article. At the 'image-flow' and 'page-flow' level these

two different elements are simultaneously playing together. The voice-over was another common strategy, adopted by five of the videos, where the visual content shown would be social media posts or recorded live-streamed footage. Lastly, there was one instance in the dataset of a news report structure, in this case, a news reporter was reporting live outside the Notre Dame Cathedral, thus mimicking a mainstream news format. A further discussion of the visual depiction of the videos will be provided in the following sections. The 15 videos selected with a brief description of each video and its macro-genre, is provided in Table 6-1.

#	Structure	Purpose of Video	Video Title
1	Vlog with social media content	Encouraging hate speech towards Muslims	The Notre Dame Fire
2	Vlog with social media content	Encouraging hate speech towards Muslims	What They're NOT Telling You About The Notre Dame Fire
3	Vlog with social media content	Encouraging hate speech towards Muslims	"We wish more fire upon you" – Muslim world reacts to Notre Dame tragedy
4	Voice-over with social media content	Explaining an arson conspiracy	Mystery figure at Notre Dame cathedral fire
5	News report with reporter outside Notre Dame Cathedral	Reporting conspiratorial discourse about French migrants	Notre Dame Cathedral Fire: Suspicion After Hundreds of French Churches Vandalised | Martina Markota
6	Vlog with social media content, dual screen	Explaining a conspiracy about Macron	Fire At Notre Dame Follows Wave Of Church ATTACKS – Will They Blame The Yellow Vest Movement?
7	Vlog with social media content	Explaining a conspiracy linked to Globalism	Notre Dame Fire: Globalist False Flag to Trigger WWII?
8	Vlog with social media content, dual screen	Explaining a conspiracy linked to Notre Dame art	Notre Dame Paris Fire INSIDE JOB, Destroying Tartarian Art by Burning
9	Voice-over with multi-media content	Explaining a conspiracy linked to the New World Order	The Notre Dame Cathedral Fire | A Planned, Deliberate Event

10	Voice-over with live-stream video content	Explaining a UFO Conspiracy	Now 3 UFOs Filmed At Notre Dame Fire! OMG!
11	Voice-over with live-stream video bond	Showing evidence of an arson conspiracy	NOTRE DAME FIRE: Suspicious Activity on Roof
12	Voice-over with multi-media content	Explaining a conspiracy linked to Freemasons	THE SCARY TRUTH ABIOUT THE NOTRE DAME FIRE THAT NO ONE IS TALKING ABOUT...
13	Vlog	Explaining a conspiracy linked to authorities	Breaking: "Nostradamus Predicted Paris Notre Dame Would Burn" / WW II
14	Vlog	Explaining a conspiracy linked to authorities	The Notre Dame Cathedral Fire was Arson
15	Vlog	Explaining a conspiracy linked to authorities	What's the truth abou the Notre Dame Cathedral Fire?

Table 6-1 – Notre Dame Fire Macro-Genre Video Dataset

6.3. Analysis of YouTube Videos: Affiliation and Legitimation

This section will detail the affiliation and legitimation analysis undertaken on the YouTube video transcripts and visual content. Whilst both linguistic and visual modes are explored, there is a shared focus on how affiliation and legitimation are enacted by both modes together. This chapter will be structured according to the key legitimation strategies uncovered across the dataset. These key legitimation strategies will be further explored according to the social bonds that were legitimated due to these strategies. The analysis refers to frequencies of particular legitimation and affiliation strategies that are summarised in Table 6-3 and Table 6-4.

As an initial insight into the data, it is important to comment on the role of the Notre Dame Fire cathedral as a bonding icon. In the case of the Notre Dame Fire, white supremacist and conspiratorial communities (communitas), rally around the shared values (doxa) of a traditional European culture (Catholic and white) that is now

threatened by a multicultural and pro-immigrant France. The role of the Notre Dame Cathedral as a bonding icon will be discussed in the following results section.

6.3.1. Technological Authority

The first legitimation strategy that will be discussed is TECHNOLOGICAL AUTHORITY. TECHNOLOGICAL AUTHORITY, a legitimation feature arising from this research, describes how the use of social media or other forms of digital media is used to construct credibility. This is an extension to the original legitimation framework that considers IMPERSONAL AUTHORIZATION as laws, rules and regulations, in other words:

> "...the answer to the unspoken 'why' question is then, not 'because I say so'.... but 'because the laws (the rules, the policies, the guidelines, etc.) say so" (Van Leeuwen, 2007: 96).

This thesis extends IMPERSONAL AUTHORIZATION to consider the answer to the 'why' question to be *because social/digital media said so*.

Referring to screenshots or videos was a frequently used strategy by the YouTubers to provide evidence supporting their opinions. Screenshots, also known as screen captures or screen grabs, are still images that replicate a computer screen or the screen of a mobile device. They are often recontextualised, for instance, by being pasted inside another text such as a social media post. Screenshots have become a ubiquitous part of our social media behaviours. Screenshots are associated with specific practices and values in society, such as providing substantiation of online activity, as a kind of documentary archive of content to peruse later, and as a tool to make personal encounters public. Since screenshots are so closely aligned with these values of substantiation and archiving, they are often believed to be true copies of content (Jaynes, 2020), factoring out the possibility that they are mock-ups or manipulated reproductions. This can be dangerous in terms of fuelling misinformation and disinformation practices.

In the following example, the YouTuber verbally reacts to footage of people laugh reacting (using an emoji as a token of laughter) to news of the Notre Dame Cathedral on fire, in order to construct an 'evil Muslims bond':

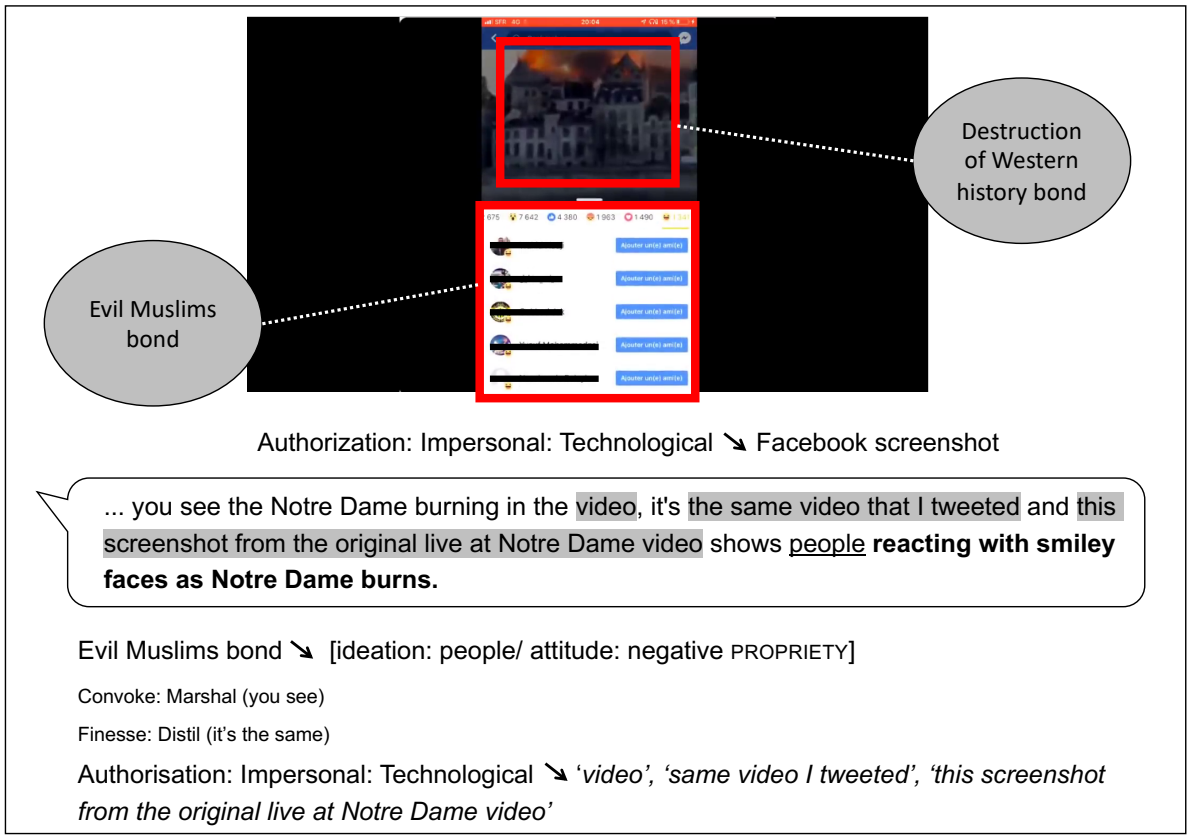


Figure 6-1 – Screenshots of Social Media Posts from the Notre Dame Fire Dataset

This is an image of a video on Facebook with some of the unfolding feed of reactions visible underneath. This screenshot was used in three of the videos from the Notre Dame Fire dataset to deceptively claim that Muslims are 'evil' and unable to integrate into western society because they have no respect for Western culture and traditions. This sort of screen recording was used by YouTubers to promote white supremacist values about the superiority of western society and to assert that immigrants are threatening the survival of western culture.

Within this kind of white supremacist ideological framework, the salient image of the Notre Dame Cathedral on fire is a 'bonding icon' (as illustrated in Figure 6-2) (Stenglin, 2008; Tann, 2012) and represents western history and culture (the Cathedral) being

destroyed (by fire). The other salient visual element in the screenshot is the depiction of Muslims insulting Western society by 'laugh reacting' to the Notre Dame Cathedral on fire (the YouTuber states the users are Muslim due to their Arabic-sounding names) that counteracts the bonding icon. Afterwards, a screen recording is shown that scrolls through a list of laugh reactions to show a list of Arabic-sounding names that the YouTuber associates with Muslims.

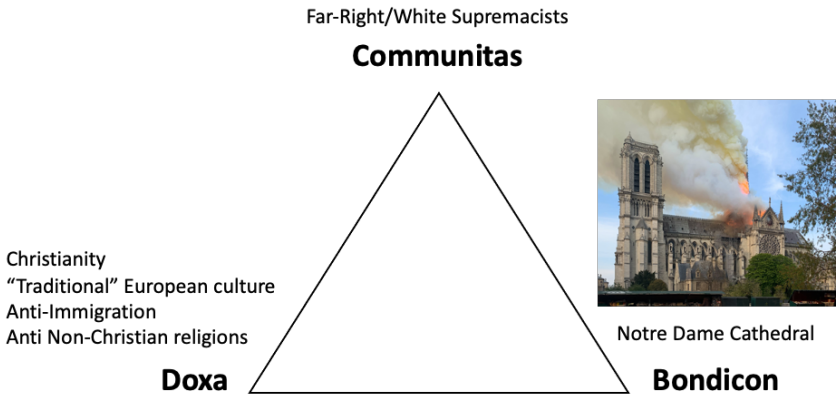


Figure 6-2 – Notre Dame Cathedral as a Bonding Icon (adapted from Tann, 2011)

The scrolling video provides a sense of dynamic unfolding that also contributes to TECHNOLOGICAL LEGITIMATION, creating the sense that the YouTuber is showing evidence in real-time when promoting the view that Muslims do not align with Western society. The negative PROPRIETY in the voice-over aligns with the values depicted in the visual content: '*people*' (which the YouTuber later in the video identifies as Muslims) are described as '*reacting with smiley faces as Notre Dame burns*'. This negative evaluation targeted at Muslims forms an 'Evil Muslims bond' that resonates throughout the video. As the annotation in Figure *6-1* suggests, legitimation is also occurring verbally, in the identification of particular technological semiotic entities such as '*the same video I tweeted*' and '*this screenshot from the original live at Notre Dame video*'. These entities highlight the importance of technology, namely social media posts and screenshots, in supporting the YouTuber's accusation that Muslims were reacting with impropriety to the Notre Dame Cathedral on fire. Classifying the video as '*original live*' also emphasises the legitimacy of the screenshot through positive APPRECIATION invoking VERACITY. Thus, in this example we see an alignment between the values depicted verbally and visually,

and a coordination of visual and verbal legitimation strategies cantering on TECHNOLOGICAL AUTHORITY.

TECHNOLOGICAL AUTHORITY was also used as a strategy to justify conspiratorial discourse about authorities. In this example, by showing a live stream from ITV News, the YouTuber uses what they portray as first-hand evidence to support the YouTuber's own claims that the fire is suspicious:

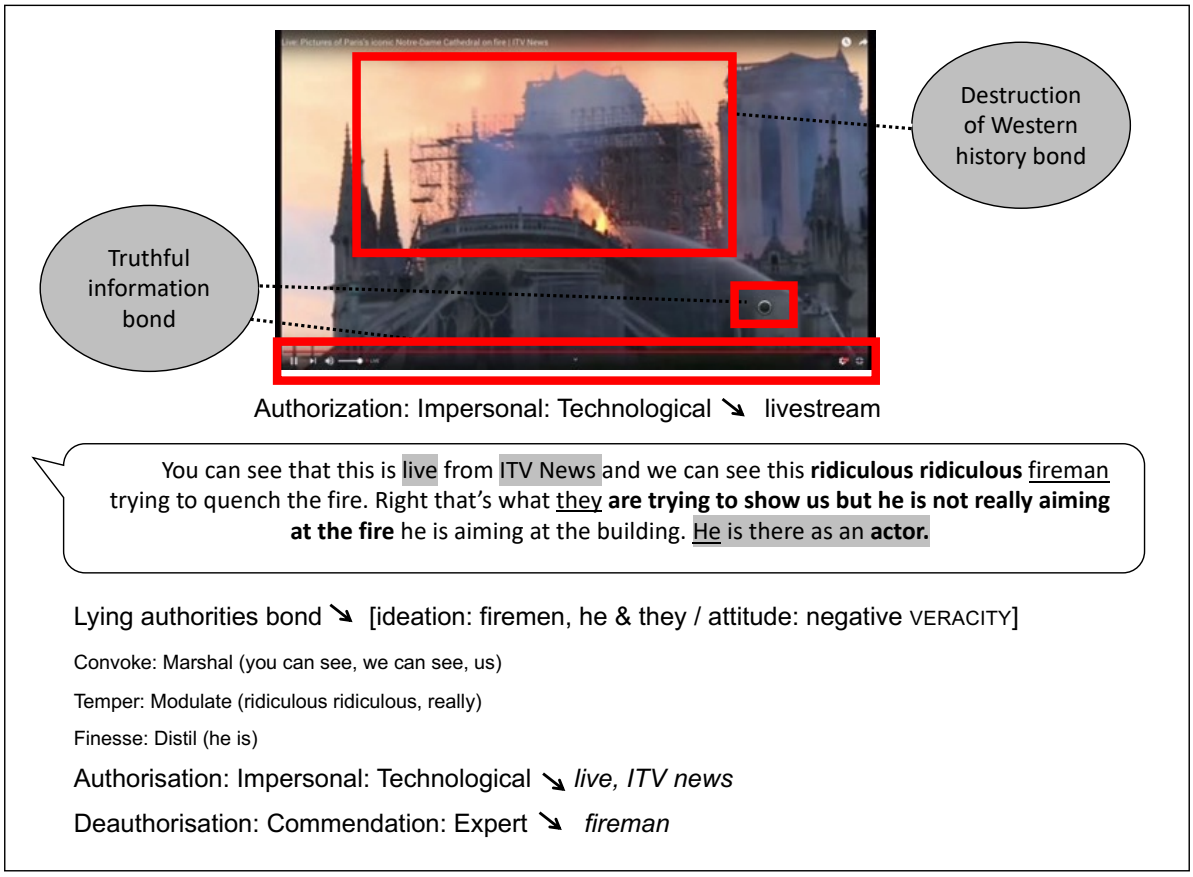


Figure 6-3 – Screenshot of a Live TV News Stream

The fireman is negatively evaluated for pretending that he is quenching the fire. ITV News does not promote this claim, rather the YouTuber has formed their own interpretation of the video, despite playing the same video as ITV News. The CONVOKING affiliation of '*you*' and '*us*', like the previous example, addresses these evaluations directly to the audience and emphasizes the TECHNOLOGICAL AUTHORITY of a live video. The YouTuber has also negatively evaluated ITV news and 'authorities', who are

referenced as '*they*', for setting up this scenario and using an '*actor*'. The phrase "*they are trying to show us*" invokes this negative assessment as a type of engagement resource that claims they have set this up. At the end of the sentence, the DISTILLING affiliation ("is there") definitively states that the fireman is there as an actor, creating an additional 'lying authorities bond'. Again, what is interesting about this example, is how a video is played from a mainstream media outlet that has not been edited, and yet the YouTuber has changed the meaning of the video via their evaluative language and the affiliation strategies they use in order to get users to rally around their claims. These two examples show how TECHNOLOGICAL AUTHORITY is an important tool for a YouTuber to convince an audience that their claims are truthful and get audiences to rally around the key bonds they have promoted.

Screenshots can also make reference to other technologies such as mobile messaging platforms in order to invoke addition layers of TECHNOLOGICAL AUTHORITY that aid in the construction of legitimacy. In the Notre Dame Fire corpus there were instances where screenshots and screen recordings of live footage of the Notre Dame on fire were employed by YouTubers as evidence that 'drones' or 'UFOs' were involved in the fire. In one example a YouTuber took a screenshot of a recording of the Notre Dame Fire that was embedded in a tweet, they then recorded themselves repeatedly zooming in on a black speck in the image, which they identify as a drone (Figure 6-4). In this instance, the tr-3b (drone) is a positive invoked evaluation because it is deemed as real ("*the photos done and it had a tr-3b in it*"), establishing a 'real drone (tr-3b) bond'. In terms of legitimation strategies, the presence of the drone (tr-3b) is a form of TECHNOLOGICAL AUTHORITY – the YouTuber's evidence that the fire was deliberate. Additionally, the '*twitter account*' from which the footage was taken is also an example of TECHNOLOGICAL AUTHORITY because it is depicted as evidence of the tr-3b drone appearing before the Notre Dame Cathedral was on fire.

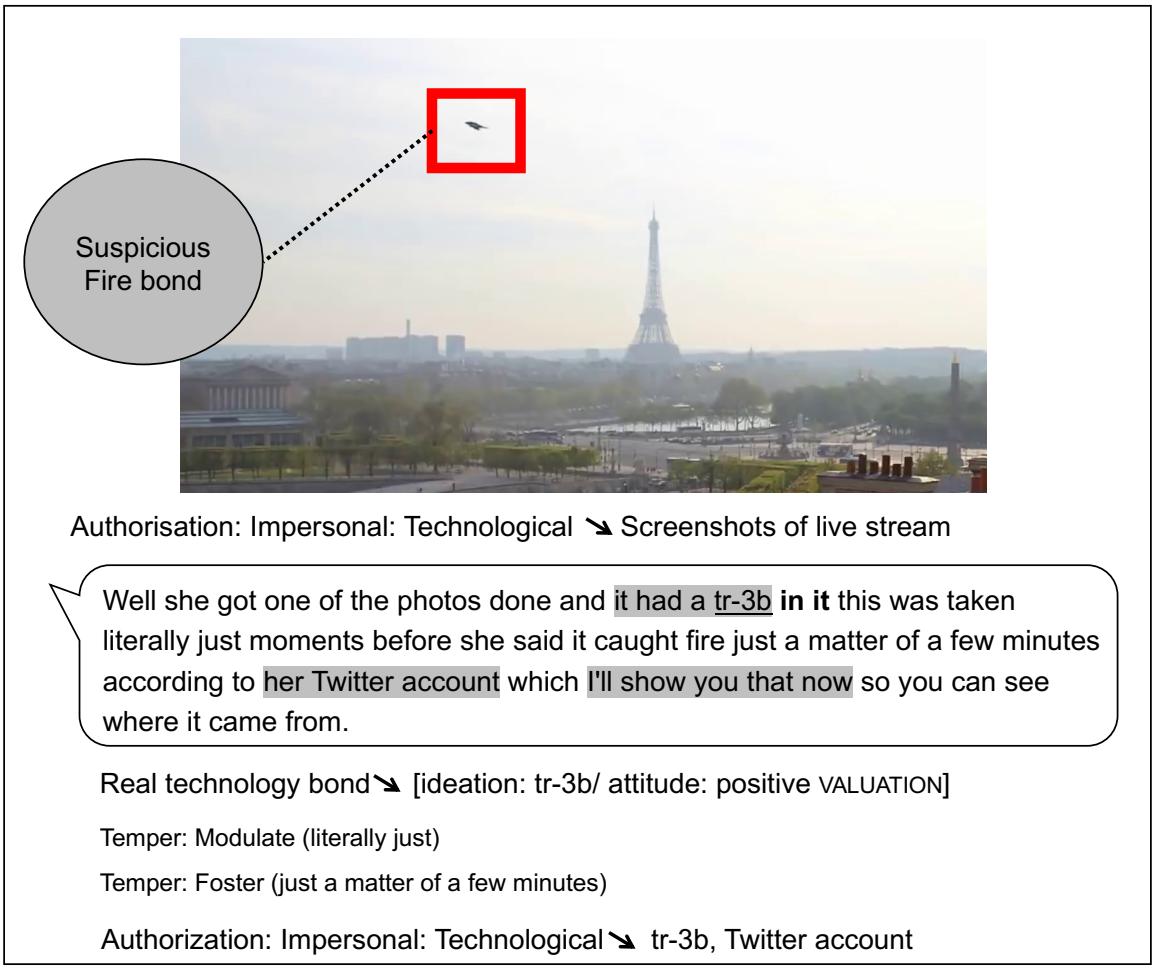


Figure 6-4 – Screenshot Referencing Other Technologies from the Notre Dame Fire Dataset

Screenshots of online articles, as a source of information, were also used as evidence to support deceptive and conspiratorial claims. For instance, the following example featured an article from the Daily Star website, with the salient headline '*Mystery 200-year-old letter revealed World War 3 plans – and final battle against Islam*' (Figure 6-5). This headline promotes an 'evil Islam bond' with the phrase the '*final battle against Islam'* invoking negative PROPRIETY. Additionally, the article featuring the Daily Star logo and the structure of a news website adds a further dimension of TECHNOLOGICAL AUTHORITY by co-opting the authority of traditional media. Whilst, in the voiceover, the paper is initially negatively evaluated as something the YouTuber would not usually buy, the story that the paper tells is evaluated positively, contributing to a 'truthful story bond' that supports the co-occurring 'evil Islam' bond. This is an example of how online

mainstream media is manipulated to support conspiratorial claims. In all these examples TECHNOLOGICAL AUTHORITY has a key role in legitimating suspicions about the Notre Dame fire.

Authorisation: Impersonal: Technological↘ Daily Star screenshot

It's **not a kind of paper that I would buy** but it **says here** the story, **mystery** 200 year old letter **revealed World War three plans** and **final battle** against Islam…

Unpopular paper bond↘ [ideation: it's/attitude: negative REACTION]

Finesse: Embellish (not a kind)

Truthful story bond ↘ [ideation: it & the story/ attitude: positive VALUATION]

Finesse: Distil (it says here)

Truthful letter bond ↘ [ideation: 200 year old letter/ attitude: positive VALUATION]

Evil Islam bond↘ [ideation: Islam/ attitude: negative PROPRIETY]

Authorization: Impersonal: Technological ↘ *'it says here'*

Figure 6-5 – Online Articles from the Notre Dame Fire Dataset

6.3.2. Storytelling as Moral Justification

MYTHOPOESIS, legitimation constructed via storytelling (Van Leeuwen, 2007), was used in the dataset to construct stories that provided strong moral justification for hatred towards certain religious groups such as Islam. MYTHOPOESIS can be expressed via several different storytelling techniques, such as *moral tales* (where protagonists are rewarded for their actions), *cautionary tales* (where protagonists are punished for their actions), *inversion* (using techniques such as personification and metaphor, rather than providing information in a straightforward way), *single determination* (where the story or protagonist has one purpose only) and *symbolisation* (symbolic actions representing broader values). In this dataset, MYTHOPOESIS was used to typically tell moral tales and cautionary tales about the past or future and provide evidence of why a certain moral stance was taken. For example, in the extract in Figure 6-6 from a YouTube video promoting a globalist conspiracy theory about the Notre Dame Fire, a story is told about the meaning behind each world war, constructing a 'powerful wars bond':

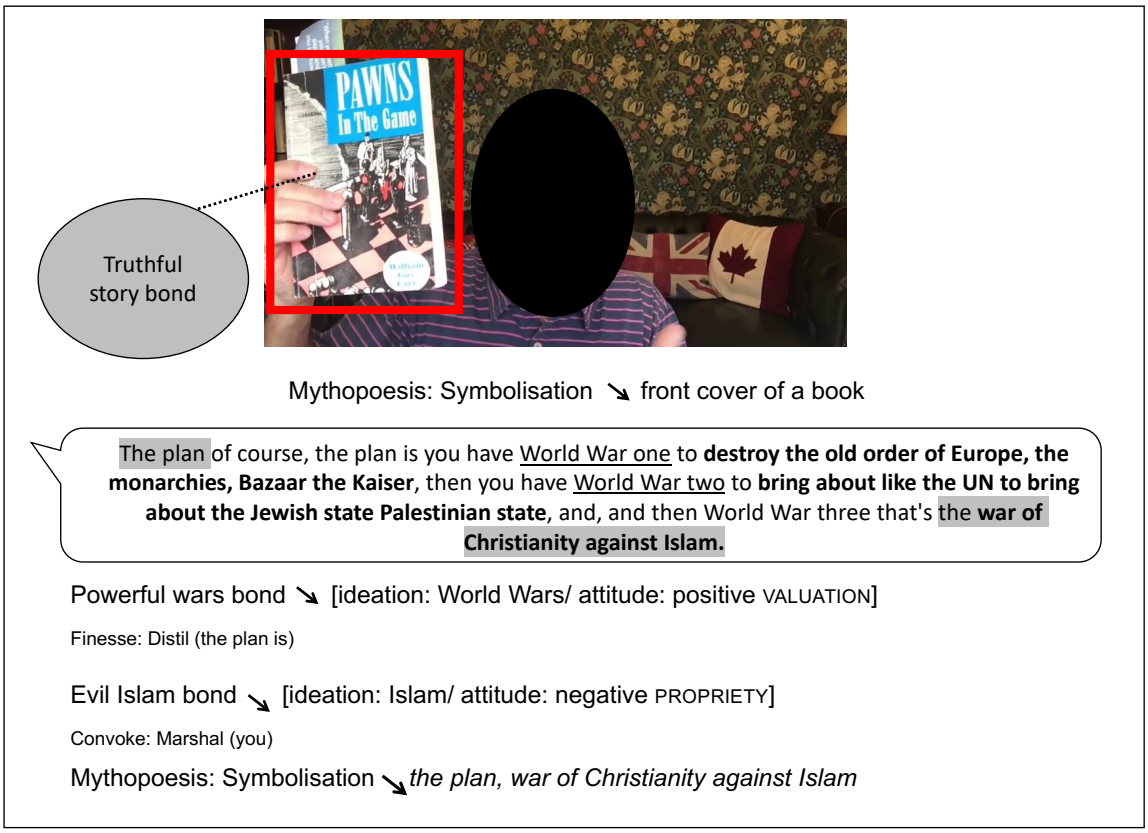
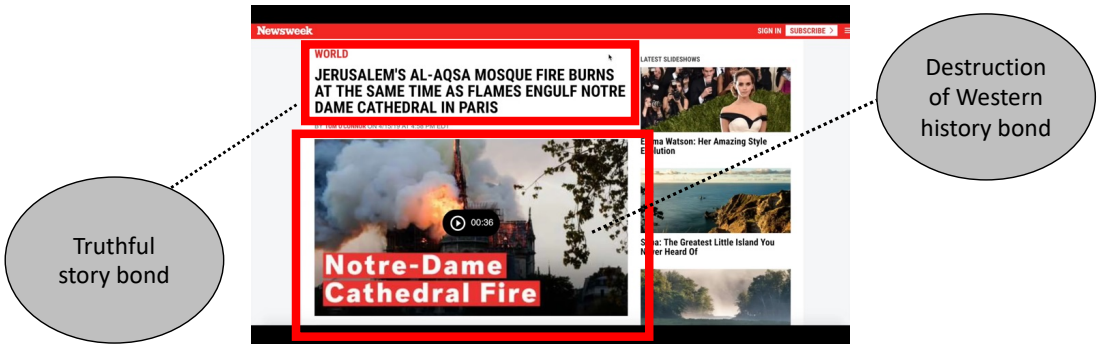


Figure 6-6 - Image of a YouTuber Vlogger

In this example, there is repeated reference to '*a plan*' and a synopsis of world wars. This constructs a brief explanation of the plan, connecting events together. '*World wars*' has POSITIVE VALUATION because these wars have brought about outcomes, so in other words, they are powerful. There is a lack of inscribed evaluation in this example, but Islam is still evaluated as evil at the end, with the DISTILLING affiliation strengthening the claims throughout this text by providing no other alternatives. By using these storytelling and symbolisation techniques, the YouTuber constructs their moral legitimation to explanation why the West is currently threatened.

SINGLE DETERMINATION was also an important strategy in videos that emphasised conspiratorial thinking. For example, in the following excerpt, the phrase '*this all connects' is* a legitimation strategy by the YouTuber to justify their conspiratorial thinking:



Authorization: Impersonal: Technological ↘ online article

Yeah **funny** how this how this all connects and you know, part of the plan, look I've talked, I talked about this, the plan, the **big** plan is to **erase everything** and that that includes Islam as well, it's not just Christianity.

Valuable conspiracy bond↘ [ideation: this (conspiracy)/ attitude: positive VALUATION]

Temper: Modulate (all, big)

Bad and powerful plan bond ↘ [ideation: plan/ attitude: positive VALUATION]

Powerless Islam and Christianity bond ↘ [ideation: Islam/ attitude: negative VALUATION]

Finesse: Distil (not just)

Mythopoesis: Single Determination↘ *this all connects*

Figure 6-7 - Screenshot of an Online News Article

'*This*' refers to the YouTuber's conspiracy that there is a plan to erase all religions. The repeated ideational target of '*the plan*' and how '*this all connects*', is a part of the YouTuber's MYTHOPOESIS legitimation strategy, by telling a story with a SINGLE DETERMINATION – that everything can be linked together. The TEMPERING affiliation of '*all*', '*big*' and '*everything*' throughout the excerpt emphasises the powerfulness of the plan. So whilst the plan is evaluated negatively and as powerful, Islam and Christianity are contrasted as powerless. The TEMPERING affiliation of '*not just*' again emphasises the point that both Islam and Christianity are powerless. Overall, storytelling acts as a moral justification for the YouTuber's stance to invite the audience to engage in conspiratorial thinking.

6.3.3. Rationalisation

RATIONALISATION is concerned with how knowledge is constructed (the rational sphere) and is the opposite to the legitimation strategy of MORALISATION, that is instead concerned with emotional judgements (Van Leeuwen, 2007: 100). The repetition of engagement strategies like '*provable*' throughout the video transcripts indicated the YouTubers' concern with constructing a sense of legitimacy about their claims by appealing to the epistemic authority of the domain of scientific research. Although we might associate conspiracists with MORALISATION (arguments surrounding emotions and subjective ethical viewpoints) rather than RATIONALISATION, the transcripts all expressed this concern with using research to discover the truth. Nonetheless, this research was defined by YouTubers as a personal investigation rather than institutionalised research such as that associated with educational institutions. The importance of conducting your "*own research*" was emphasised by YouTubers with positive evaluation, appealing to people's sense of rationalisation and desire for first-hand knowledge. This is shown in the following example:

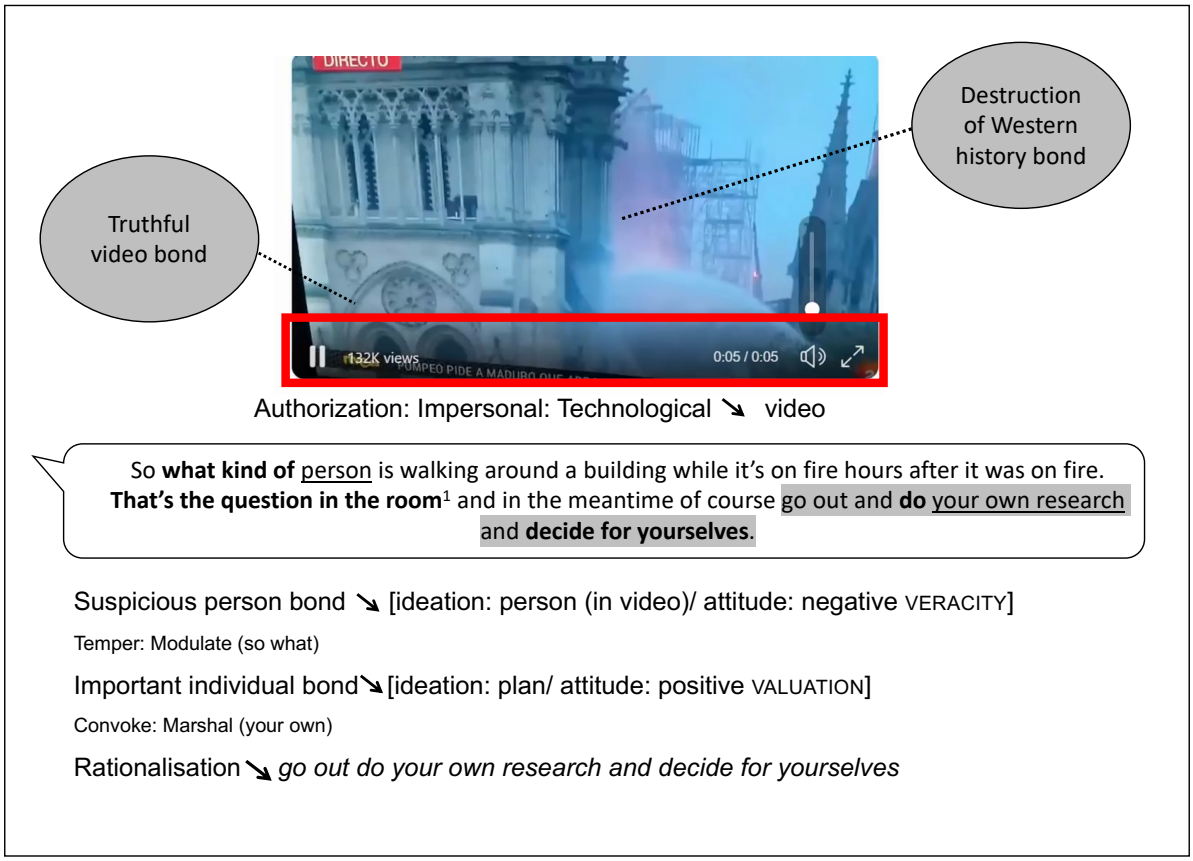


Figure 6-8 - Screenshot of Looped Live TV News Footage

The YouTuber firstly evaluates the person in the video with invoked negative evaluation for their suspicious behaviour ("*so what kind of person*"), emphasising this suspicious behaviour with the TEMPERING affiliation resource of '*so what*'. The next coupling, '*do your own research*' can be associated with POSITIVE VALUATION, forming an 'important individual bond', in the sense that it is important for the individual to do their own research. The particular phrase '*your own*' is an example of CONVOKING AFFILIATION because it enhances the 'important individual bond' by directing these thoughts to the audience. The suggestion to '*do your own research*' is also a RATIONALISATION strategy which legitimates first-hand knowledge by again emphasising the individual's autonomy. A common pattern among the transcripts involved the YouTuber firstly evaluating a video, providing their own opinion, but then resorting back to RATIONALISATION strategies to emphasise the importance of free agency, in other words, an individual '*doing their own research*'.

In the dataset, there were examples of 'evidence collages' – collections of multiple screenshots displayed simultaneously on the screen. According to Krafft and Donovan (2020: 205) an evidence collage is formed through 'image files that aggregate positive evidence'. This positive evidence constitutes 'direct proof' of the issue/fact that is at stake, in contrast with circumstantial proof which refers to evidence not drawn from direct observation. In the Notre Dame Fire dataset, evidence collages typically consisted of multiple tweets shown side by side that functioned to add an extra sense of validity to the claim being made. For example, in Figure 6-9 two tweets are shown side by side that both state churches have been attacked and which both include images of damaged churches. These images of churches function as bonding icons because they represent Catholic culture rather than just a building, evoking negative interpersonal meanings about the destruction of French culture and society contributing to a 'Destruction of Western History bond'. The reference to '*multicultural France'* links this to anti-immigration stances. By displaying multiple screenshots, the YouTuber is trying to construct a greater sense of legitimacy by highlighting that multiple people have shared this information so it must be true. The phrase '*I'll leave links*' by the accompanying voiceover also reinforces TECHNOLOGICAL AUTHORITY, with the implication that, because there are more than one link and the constitute '*information*', they are a legitimate source and part of the 'truthful information bond' established. This also serves as a RATIONALISATION strategy in terms of legitimation, solidified by the instruction to '*double check it'* that appeals to people's sense of rationality.

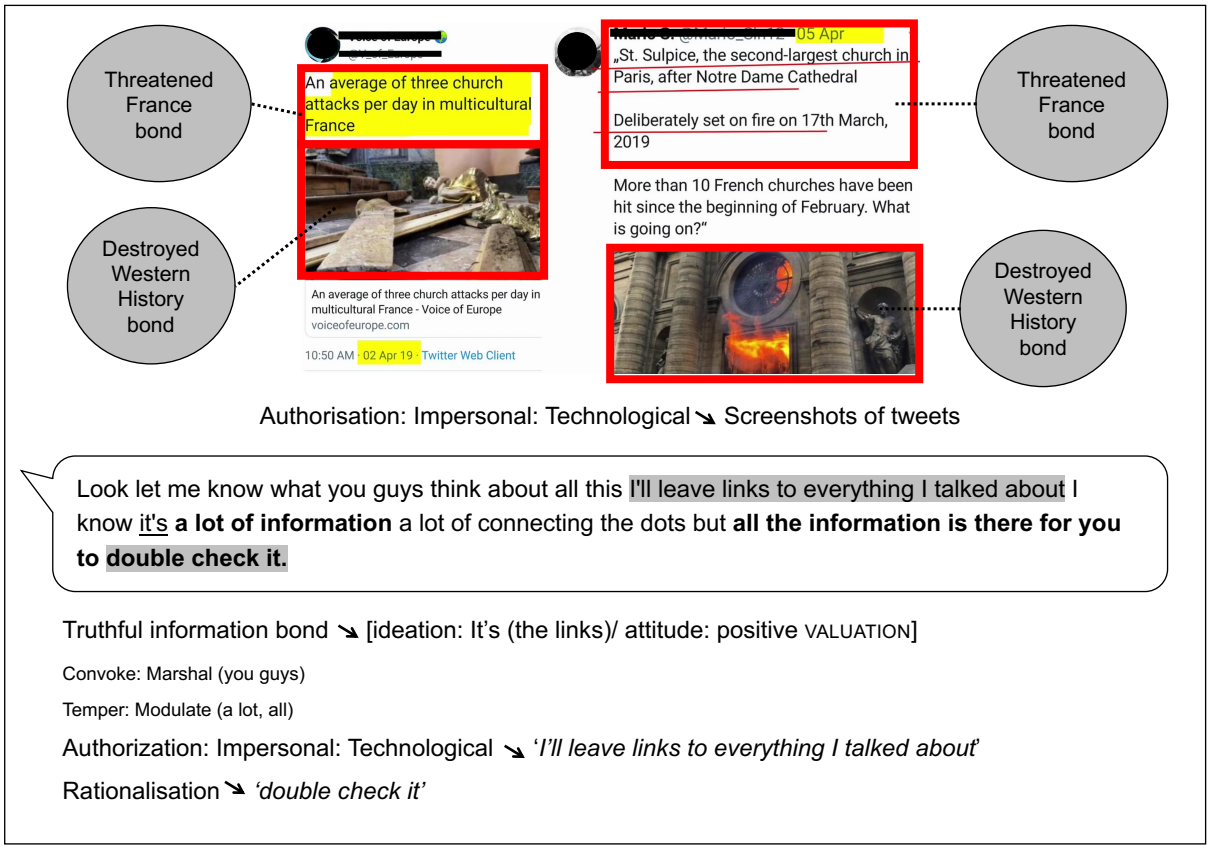
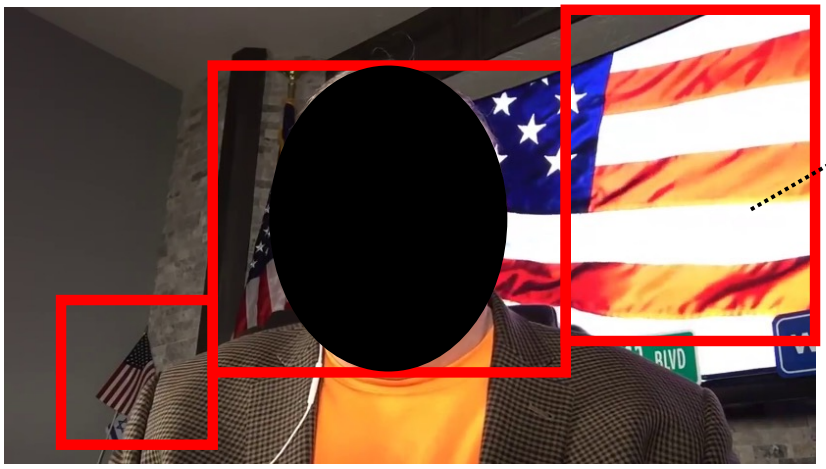


Figure 6-9 - Evidence Collage from the Notre Dame Fire Dataset

Across the entire dataset, three out of the 15 videos did not feature visual screenshots and instead invoked the presence of screens in the linguistic verbiage. These three videos employed the vlog macro-genre (a YouTuber directly speaking to the camera). A example of this is illustrated in Figure 6-10. The YouTuber refers to content present on a screen (an article discussing churches being destroyed across all of France) as a form of TECHNOLOGICAL AUTHORITY. They express negative REACTION to the fire, contributing to a 'Suspicious Fire bond'. In terms of legitimation strategies, the YouTuber refers to an 'article' and, whilst there is no corresponding visual content, they are simultaneously looking at a screen (presumably containing the article).



Authorisation: Personal: Individual ↘ YouTuber

So guys this is an article that came out in March 21st of this year just March 21st we're talking 25 days before the Church of Notre Dame burns aground and you're immediately going to say oh no it wasn't arson it wasn't terrorist attack oh no no no guys **let's get real here** you've got to investigate…

Suspicious fire bond ↘ [ideation: it (fire)/ attitude: negative REACTION]
Convoke: Marshal (guys, you're)
Temper: Modulate (oh no no no)

Authorisation: Impersonal: Technological ↘ an article
Rationalisation ↘ '*investigate*'

Figure 6-10 - YouTube Vlogger

Overall, it is important to consider how conspiratorial YouTube videos use RATIONALISATION as a technique to persuade audiences that their information is correct and replicable. Additionally, it is important to consider where these RATIONALISATION strategies are placed in relation to the entire video. For example, if the RATIONALISATION strategy is located at the end of the video, the YouTuber has already had time to set up certain bonds that will influence the viewer's notions of research and truth. This is why the examples shown typically have RATIONALISATION strategies at the end of their videos, after firstly using strategies like TECHNOLOGICAL AUTHORITY to present initial information. In all, these examples show how RATIONALISATION was an important tool used in order to persuade audiences of a video's credibility.

6.3.4. Evaluating Religions

MORAL EVALUATION was an important legitimation strategy for evaluating religions because it appealed to viewer's personal belief systems, therefore allowing for more inscribed evaluations. These examples of inscribed evaluation were mainly directed towards comparing Christianity versus Islam, and emphasising an individual's own ethical stance rather than being "imposed by some kind of authority without further justification" (Van Leeuwen, 2007: 97). In the following example, the YouTuber is explaining to the audience why Christians are being threatened by Islam. A 'valuable Christianity' and 'righteous Christianity' bond is constructed in comparison to an 'evil Islam bond':

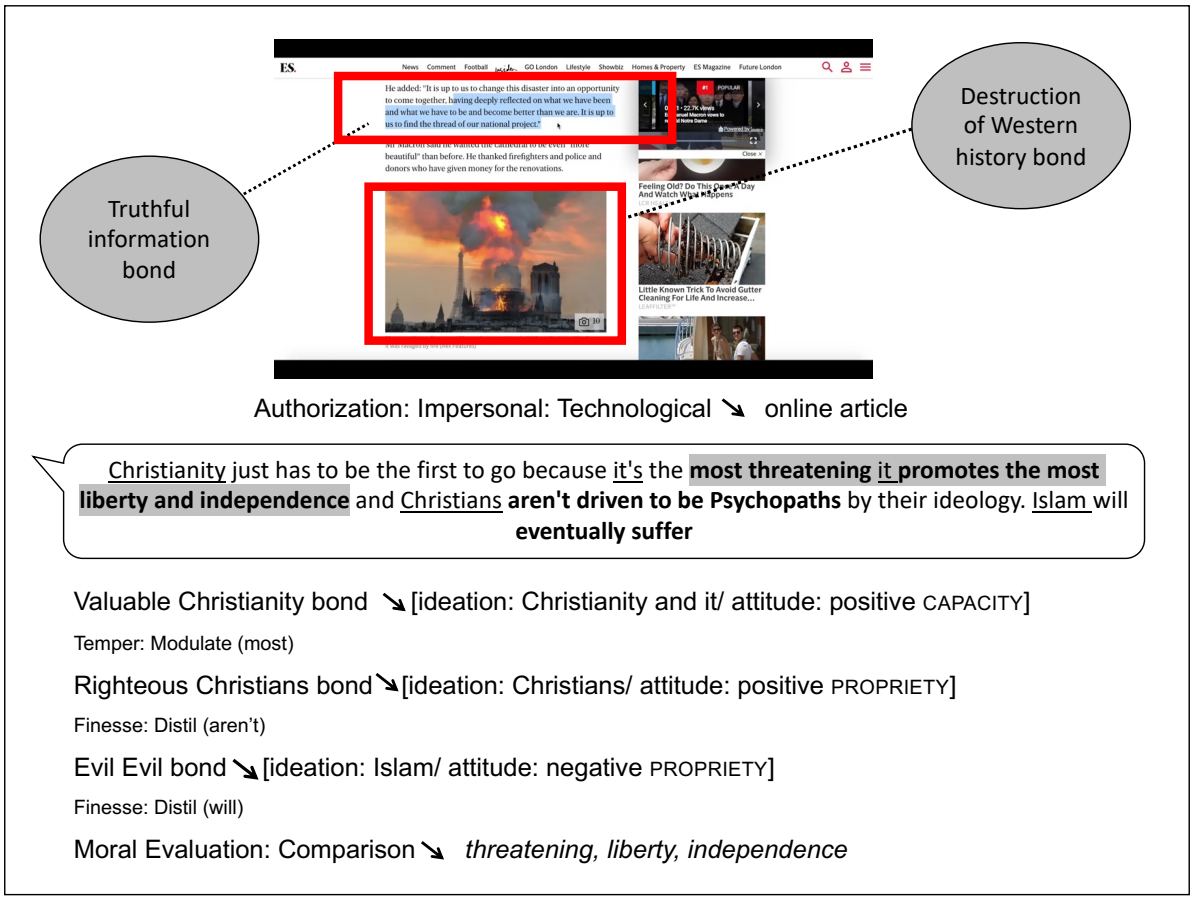
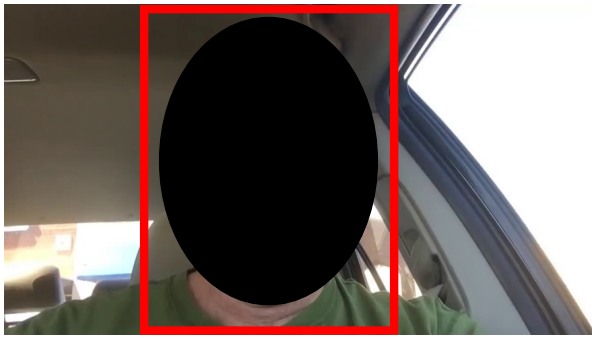


Figure 6-11 - Screenshot of Online News Article

Here, we are presented with more explicit evaluation: Christianity is positively evaluated because it is described as threatening, which is emphasised with the

TEMPERING affiliation of '*most threatening*' (an example of tempering because it is a superlative) and associated with positive values such as liberty and independence to form a 'valuable Christianity bond'. Christians are also positively evaluated, with the FINESSING affiliation of '*aren't*' definitively situating Christians as 'righteous'. In contrast, Islam is negatively evaluated because it is associated with psychopathic behaviour. The FINESSING affiliation of '*will*' definitively proclaims an 'evil Islam bond'. This comparison of Christianity and Islam, constructed by different polarities of evaluation, is also a MORAL EVALUATION legitimation strategy due to its focus on constructing righteous versus evil bonds.

Moral evaluation was used as a strategy to convince the audience to affiliate with Christianity. In the next example, a warning is given to those who do not turn back to Christianity:



Authorisation: Personal: Individual ↘ YouTuber

I'm going to say something provocative, this is a **minor chastisement**, it's a tiny taste of the **horrors coming, if the West does not turn back** to Jesus Christ.

Tragic event bond ↘ [ideation: this & it's/ attitude: negative VALUATION]

Finesse: Distil (it's)

Important Jesus Christ bond ↘ [ideation: Christians/ attitude: positive PROPRIETY]

Convoke: Designate (the West)

Moral Evaluation: Comparison ↘ *horrors, turn back*

Figure 6-12 - YouTube Vlogger

Rather than appeal to the church as an institution, the YouTuber appeals to an audience that is frightened about further horrors after the Notre Dame fire. The pronouns '*this*'

and '*it*' reference the Notre Dame Fire that is associated with NEGATIVE VALUATION. The definitive language associated with these pronouns (e.g., "*this is a minor chastisement…*", "*it's a tiny taste of the horrors coming…*") and negative evaluative language forms a tragic event bond. In contrast, '*Jesus Christ*' is evaluated with POSITIVE PROPRIETY for being a beacon of hope. CONVOKING affiliation directs this comment to Westerners, forming an 'important Jesus Christ bond'. It is the use of evaluation in this comment that legitimises moralistic feelings. Overall, the purpose of MORAL EVALUATION in this dataset was to convince audiences to rekindle their Christian faith and fight against non-Christians.

6.3.5. Deauthorising Authorities

Throughout the dataset, delegitimation strategies that targeted 'authorities' such as politicians and public officials were used to support negative evaluations of these authorities and rally audiences in their hatred of authorities. DEAUTHORISING AUTHORITIES is aligned with Van Leeuwen's categorisation of COMMENDATION, that encompasses (de)legitimation occurring via expertise (expert authority) or via opinion models or role models (role model authority) (Van Leeuwen, 2007: 94-95). These DELEGITIMATION strategies worked in conjunction with negative evaluation strategies to form an 'evil authorities bond', as in the following example:

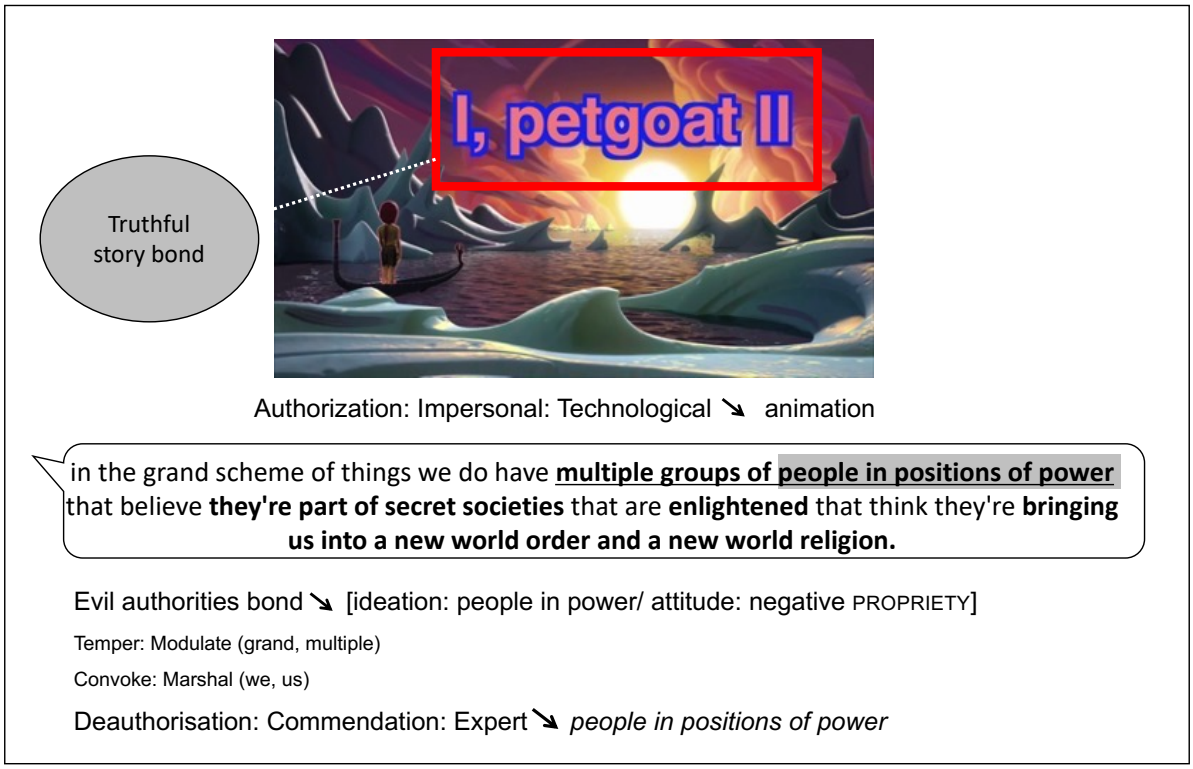
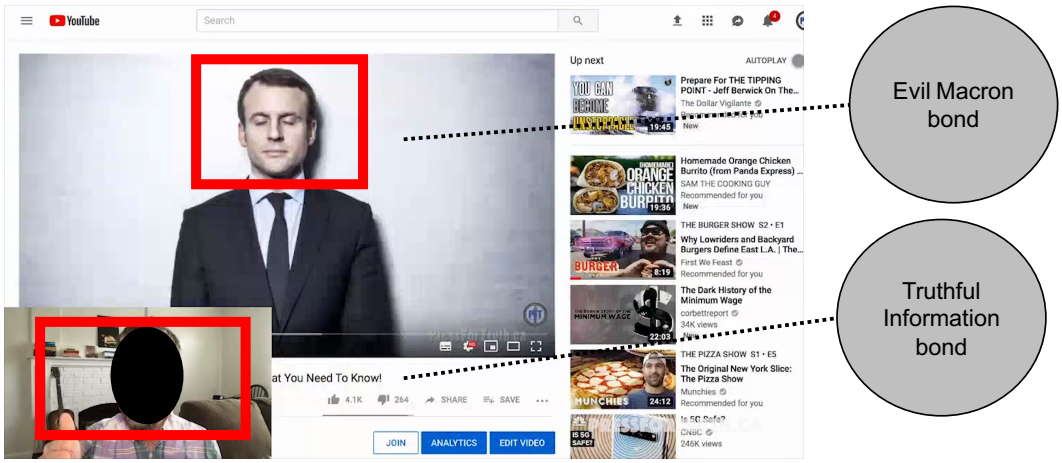


Figure 6-13 - Screenshot of an Intertextual Reference

In this example, people in positions of power are delegitimised by being negatively evaluated for being part of secret societies with suspicious activities. '*Grand*' and '*multiple*' as exaggerative adjectives are used as TEMPERING affiliation strategies to emphasise these claims whilst the vocative '*us*' as CONVOKING affiliation addresses this text directly to the audience. It is these two affiliation strategies in conjunction with negative evaluations that help to emphasise an 'evil authorities bond'. As this example shows, negative evaluations, affiliation, and delegitimation strategies work in tandem in order to deauthorise authorities.

Strategies of delegitimation were also directed towards specific politicians. As in this example, a 'lying authorities' and 'lying Macron bond' is predominately shared:



Authorisation: Impersonal: Technological ↘ Facebook screenshot

Authorisation: Personal ↘ YouTuber

If you're not familiar with what this man is like please check out my video: "the **shocking truth** about Emmanuel Macron what you need to know". Folks this guy was abused as a child by pedophiles and that's not conspiracy theory that that's that's **documented provable facts**. He then went on to **become a Rothschild banker** who went on to become the president of France, folks this man **cannot be trusted** and the fact that all these arson attacks being proven to be being **done on purpose**.

Evil Macron bond ↘[ideation: Macron/ attitude: negative PROPRIETY]
Convoke: Marshal (if you're not)
Truthful information bond ↘[ideation: that (story)/ attitude: positive VALUATION]
Finesse; Distil (documented)
Evil Macron bond ↘ [ideation: he & this man/ attitude: negative PROPRIETY]
Convoke; Marshal (folks)
Authorisation: Impersonal: Technological ↘ *'check out my video'*
(De)Authorisation: Commendation: Role Model ↘ *Macron'*
Rationalisation↘ *'documented provable facts', 'the fact that'*

Figure 6-14 – Split Screen Video from the Notre Dame Fire Corpus

Screenshots were frequently incorporated into split screen videos, adding additional layers of technological authority to the YouTuber's video. The typical structure for these sorts of videos featured one video in the corner (usually of the YouTuber) and the other video taking up most of the screen (usually a screen shot or recording), as is seen in Figure 6-14. This video incorporates a screenshot of the YouTuber's video about Macron that they are directing the viewers to listen to later. The YouTuber is present in a smaller shot in the bottom left corner in front of the screenshot.

The screenshot functions to emphasise two key bonds construed by the couplings shown in the figure, an 'evil Macron bond' realised in the heading '*The Shocking Truth*

about Macron' and a 'truthful information bond' realised by YouTube's signature platform. In the voice-over, we are presented with multiple negative evaluations targeted at Macron, contributing to the 'evil Macron bond'. The direction to '*check out my video*' supports the visual content, legitimising the screenshot as a source of evidence. The screenshot also acts as a catalyst for other negative evaluations, with the verbal voiceover also promoting these bonds. For example, through APPRECIATION invoking positive VERACITY (e.g. '*documented provable facts'*) that are used to heighten the YouTuber's apparent credibility in contrast with Macron's lack of credibility '*this man cannot be trusted*'. The choice of the vocative, '*folks'*, is an attempt to convoke the ambient audience around these values by constructing them as peers. In all, the split screen appears to have two functions: it either privileges the screenshot as an essential piece of information, or privileges the YouTuber speaking as a legitimate voice of knowledge in the video with the screenshot adopting a more evidentiary role. In all, the DELEGITIMATION of authorities was an important strategy by the YouTuber for normalising why it is fine to adopt conspiratorial thinking, as they would infer that all trust in authorities has been eroded.

Screenshots and screen recordings were often edited by the YouTuber in the Notre Dame Fire dataset, in order to cast doubt on the recordings and help align the footage with the conspiracy theory that the YouTuber was attempting to promulgate. In one example, the videos shown by the YouTuber consisted of recorded footage from a CBSN live stream that had been annotated by the YouTuber (Figure 6-15). These annotations consisted of hand drawn red arrows pointing to the Cathedral with commentary questioning visual features of the image (e.g. '*One fireman with one hose?*') that cast doubt on the reporting. The live stream is also used as a legitimation strategy, with the YouTuber stating that they are noticing details in real-time about the footage that the news organisation is not picking up on. In this example, the officers and correspondents are negatively evaluated, enacting 'lying authorities' and 'lying media' bonds. In terms of legitimation strategies, the officers and correspondents are deauthorised, based on this negative evaluation. A TECHNOLOGICAL AUTHORITY strategy is also present as the '*CBS clip'* is legitimated as evidence that these authorities and journalists show suspicious behaviour. Overall, annotated screenshots have the rhetorical function of persuading

the viewer to see the screenshot from the YouTuber's perspective, creating new meanings via the act of annotating the primary text as it unfolds.

Authorisation: Impersonal: Technological ↘ Screenshots of online articles

We can see how the officers are literally **walking like nothing happened** no urgency it's a drill in the following clips from CBS and we will see four correspondents and four employees of CBS that can report on the fire and two of them at least as far as I can tell **were on just vacation**. **Coincidence** they **were just on vacation** happen to be just in the vicinity seeing all that is happening **give me a break.**

Lying authorities bond ↘ [ideation: officers/ attitude: negative VERACITY]
Convoke: Marshal (we)
Lying media bond ↘ [ideation: correspondents/ attitude: negative VERACITY]
Finesse: Distil (it's a drill)
Lying media bond ↘ [ideation: they/ attitude: negative VERACITY]
Temper: Modulate (just)
Authorisation: Impersonal: Technological ↘ *'clips from CBS'*
(De)Authorisation ↘ *'the officers'*

Figure 6-15 – Annotated Screenshot from the Notre Dame Fire Dataset

6.3.6. Deauthorising Conformity

DEAUTHORISING CONFORMITY was the last common legitimation strategy where the YouTuber encouraged their audience to challenge mainstream opinions and instead adopt conspiratorial and fringe views. As Van Leeuwen writes, CONFORMING LEGITIMATION refers to the notion that something is legitimated because "that's what

everybody else does" or "because that's what most people do" (Van Leeuwen, 2007: 97). In this dataset, the YouTubers challenged the assumption that "everybody else is doing it, and so should you" (Van Leeuwen, 2007: 97) and instead they promoted retaliating against mainstream media and official stories. In this example, through negative evaluation the YouTuber does not conform with the 'official story'. As it reads:

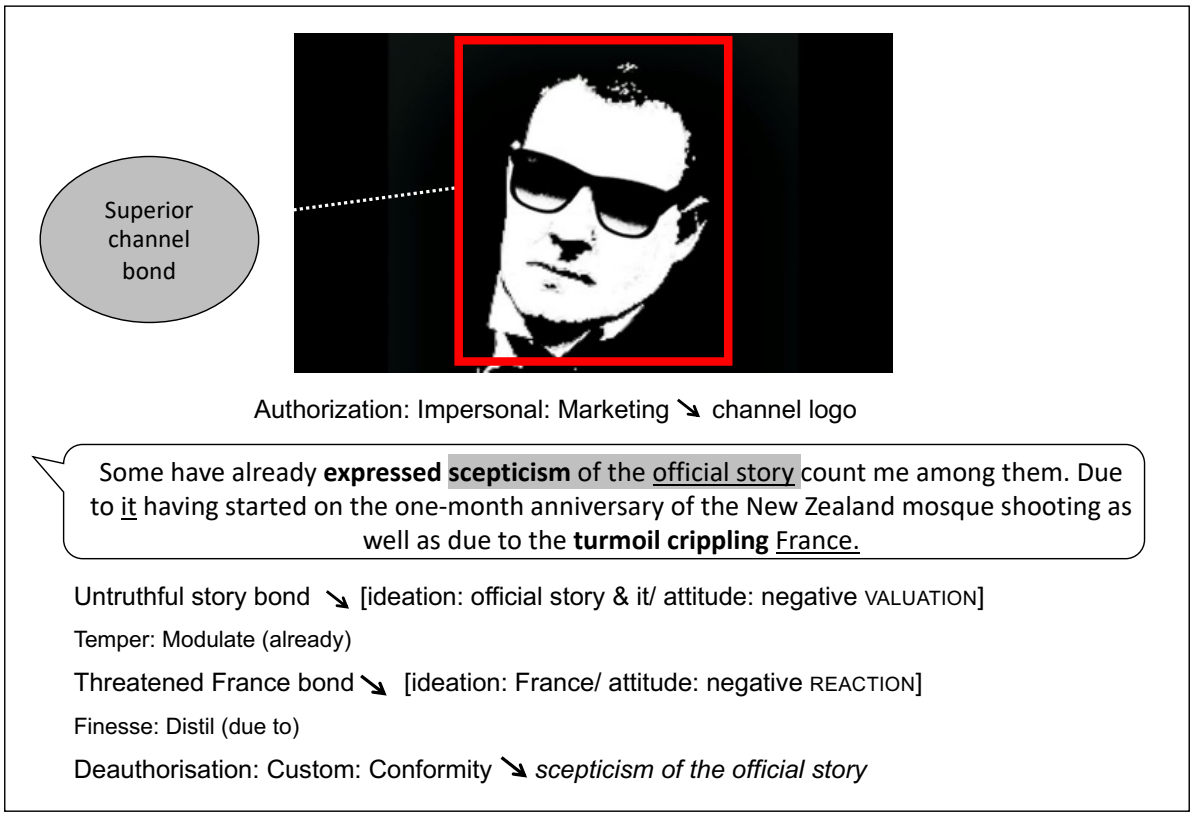


Figure 6-16 - Screenshot of YouTube Vlogger Logo

In this example, the official story is delegitimated and negatively evaluated ("*expressed scepticism*") to form an 'untruthful story bond'. So by delegitimating the official story, this means the YouTuber is not conforming with the actions of most people, who would believe the official story. France is also evaluated negatively by the YouTuber for being impacted by the Notre Dame Fire and immigration. '*Due to*' as a prepositional phrase that leaves no other alternatives is a form of DISTILLING AFFILIATION, thus definitively stating that the fire is suspicious, so this forms a 'threatened France bond' as the YouTuber makes a connection between the New Zealand Mosque shooting and France's

social and political struggles. Overall, this example encourages non-conformity and conspiratorial thinking.

DELEGITIMATING CONFORMITY and evaluating oneself as an outsider was another strategy used. For example, this YouTuber directly portrays themself as an outsider:

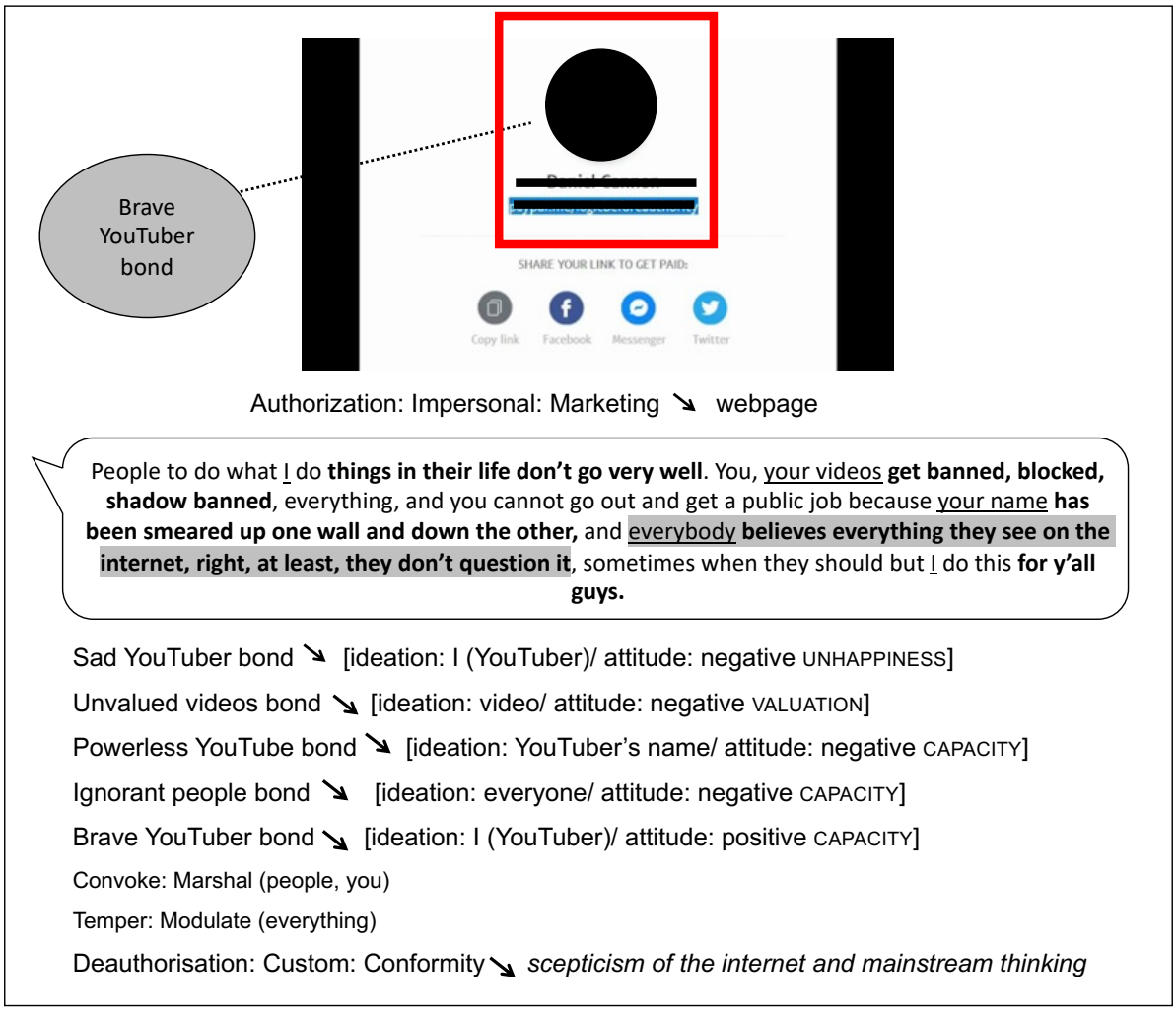


Figure 6-17 - Screenshot of YouTube Vlogger Donation Page

Here, the YouTuber is evaluating their life negatively, for having their videos banned and their inability to hold a public job. CONVOKING affiliation directly addresses this to their audience, personalising the video, and forming a 'sad YouTuber bond', 'unvalued video bond' and 'powerless YouTuber bond'. The YouTuber also negatively evaluates people who believe everything they see on the internet. TEMPERING affiliation, emphasising that '*everything*' can't be believed, promotes an 'ignorant people bond'.

Lastly, the YouTuber evaluates their ability with POSITIVE CAPACITY for taking on the task of telling 'the truth' despite personal repercussions. The CONVOKING affiliation of '*y'all guys*' again directly addresses this to the audience, emphasising the personal element of this statement and the 'brave YouTuber bond'. Across this example, the YouTuber is showing how they DELEGITIMISE CONFORMITY by living an unconventional life where they create videos that get banned and refuse to conform to accepted notions of truth. Overall, the strategy of DEAUTHORISING CONFORMITY helps to position the YouTuber as someone who will not accept official stories and mainstream media reports, instead the YouTuber portrays themself positively as someone who does not follow the crowd and questions everything. This DEAUTHORISATION strategy is particularly effective because these opinions can be easily affiliated with, as people value challenging authority and individualism, thus the YouTuber can successfully persuade the audience to share these views.

6.3.7. Summary of Results

Overall, the analysis of the transcripts and visual content revealed how affiliation and legitimation strategies work in tandem through social bonding and shared ideational targets in verbiage and visual content. The social bonds adopted by the YouTuber were used to legitimise or delegitimise people and ideas. Across the dataset, there was a wider range of legitimation strategies used compared to delegitimation strategies, with delegitimation being limited typically to authorities and conformity. The (de)legitimations strategies identified were applied throughout the dataset, whilst some videos used some (de)legitimation strategies more than others, these strategies did appear in each video. Table 6-2 summarises the two most common legitimation strategies that featured in the visual content and verbiage of each video. As this suggests, TECHNOLOGICAL AUTHORITY was the most common strategy for the majority of the videos. The outliers in this dataset as described in Section 6.2. were the vlogs that only featured the YouTuber speaking to the camera without any accompanying screenshots or footage. These vlogs tended to DEAUTHORISE authorities the most in their tactics.

Video	Most Common Legitimation Strategies (Visual)	Most Common Legitimation Strategies (Verbiage)
1	Authorisation: Impersonal: Technological Authorisation: Personal	Authorisation: Impersonal: Technological (De)Authorisation: Impersonal: Technological
2	Authorisation: Impersonal: Technological Mythopoesis: Symbolisation	Authorisation: Impersonal: Technological (De)Mythopoesis
3	Authorisation: Impersonal: Technological No other strategies	Authorisation: Impersonal: Technological Rationalisation
4	Authorisation: Impersonal: Technological Authorisation: Personal	Authorisation: Impersonal: Technological (De)Authorisation: Commendation
5	Authorisation: Impersonal: Technological Authorisation: Commendation	(De)Authorisation: Commendation Authorisation: Impersonal: Technological
6	Authorisation: Impersonal: Technological Authorisation: Personal	Authorisation: Impersonal: Technological (De)Authorisation: Commendation
7	Authorisation: Personal Authorisation: Impersonal: Technological	Authorisation: Impersonal: Technological Mythopoesis: Moral Tale
8	Authorisation: Impersonal: Technological Authorisation: Personal	Authorisation: Impersonal: Technological (De)Authorisation: Commendation
9	Authorisation: Impersonal: Technological No other strategies	(De)Authorisation: Commendation Authorisation: Impersonal: Technological
10	Authorisation: Impersonal: Technological No other strategies	Authorisation: Impersonal: Technological (De)Authorisation: Commendation
11	Authorisation: Impersonal: Technological	Authorisation: Impersonal: Technological
12	Authorisation: Impersonal: Technological Authorisation: Commendation	Authorisation: Impersonal: Technological Mythopoesis: Moral Tale
13	Authorisation: Personal No other strategies	(De)Authorisation: Commendation Authorisation: Impersonal: Technological
14	Authorisation: Personal No other strategies	(De)Authorisation: Commendation Rationalisation
15	Authorisation: Personal No other strategies	(De)Authorisation: Commendation Moral Evaluation

Table 6-2 – Summary of Legitimation Strategies in the Dataset

A summary of the types of screenshots (representing TECHNOLOGICAL AUTHORITY), categorisations and frequencies across the case study is shown in *Table 6-3*. As this table suggests, the most frequent screenshot visual structure was use of the unaltered screenshot, followed by frames that did not contain screenshots and split screen screenshots. A range of communing affiliation strategies were also employed across the entire video dataset (see Table 6-4). MARSHALLING and MODULATING were the most frequently used strategies across the video dataset, as YouTubers addressed their community as if they were talking to friends, and emphasised claims in order to make them more engaging. In contrast, DESIGNATING and FOSTERING were the least frequently used strategies, as specific communities were not frequently named and quantification was not a frequent strategy used for emphasising claims. Most of the videos used DISTILLING more than EMBELLISHING, definitively stating their arguments rather than leaving them to negotiation. Video 11 was a particular outlier in this dataset due to its short nature and lack of speaking from the YouTuber, where instead video clips were shown that the YouTuber then asked the audience to question. Overall, this analysis shows the range of communing affiliation strategies that were used in order to achieve different purposes, in regards to how social bonds were positioned.

Screenshot visual structure	Notre Dame Fire Dataset	
	Frequency	**Type**
Evidence Collage	2 (0-0.2%)	Social media
Split Screen	120 (14%)	Social media, live stream recordings
Emulated Screenshot	0 (0%)	N/A
Annotated Screenshot	38 (5%)	Social media
Unaltered Screenshot	490 (58%)	Social media, online articles
Frames that did not contain screenshots	191 (23%)	N/A
Total frames analysed	841	
Verbal reference to screenshots (number of times spoken about)	82	Social media, online articles, live streams
Total transcript words	18 306	

Table 6-3 - Frequencies and Categorizations of Screenshots in the Dataset

Video	Total Couplings Identified	MARSHAL (directly addressing a person or community)	DESIGNATE (naming a community)	MODULATE (adjusting scope of venture)	FOSTER (adjusting degree of venture)	DISTIL (making a definitive statement)	EMBELLISH (entertaining other possibilities)
1	84	28%	5%	24%	5%	31%	7%
2	50	28%	2%	22%	4%	34%	10%
3	82	48%	2%	30%	1%	16%	2%
4	48	29%	0%	35%	0%	25%	10%
5	23	26%	4%	35%	0%	4%	30%
6	97	28%	1%	29%	5%	18%	9%
7	38	29%	3%	29%	0%	13%	26%
8	90	32%	0%	39%	1%	20%	8%
9	59	44%	2%	29%	2%	15%	8%
10	28	50%	0%	32%	0%	7%	11%
11	5	0%	0%	40%	0%	20%	40%
12	65	32%	2%	42%	2%	15%	8%
13	39	38%	0%	36%	0	13%	13%
14	32	16%	0%	31%	0%	34%	19%
15	44	18%	0%	39%	5%	18%	20%

Table 6-4 - Communing Affiliation Strategies across the Transcript Dataset

6.4. Illustrative Examples from the Dataset

This section will provide detailed illustrative examples of the two key macro-genres discovered in the dataset: the social media voice-over macro-genre and the conspiracy vlog macro-genre in order to show how the affiliation and legitimation strategies work in particular macro-genres. This is in contrast to the previous section that focused on trends in affiliation and legitimation strategies across the entire dataset. This section will discuss the key features of the macro-genres, and how different legitimation strategies are used throughout the video in order to legitimatise the key social bond that the Notre Dame Fire is suspicious.

6.4.1. Social Media Voice-Over Macro-Genre

Videos in the social media voice-over macro-genre, focused on showing screenshots or live TV footage as the core visual content. While the YouTuber provides their commentary on this content, they do not show their face (in contrast to the YouTuber vlog macro-genre explained in the next section). In this macro-genre, the YouTuber aims to intrigue viewers about the suspicious nature of the fire through EMBELLISHING strategies. In contrast to DISTILLING, EMBELLISHING can open a claim to ambiguity by refusing to take a definitive stance and instead entertaining other possibilities or expressing uncertainty. Instances of EMBELLISHING were less common across the dataset compared to DISTILLING. However, EMBELLISHING was still used as a strategy to allow the audience to come to their own conclusions, as shown in these examples from two different videos (ideation underlined and attitude in bold):

1. Please pay attention. **What was** <u>that</u>? A lighter? A match? A welding flame? A tool reflecting the sun? A delayed-action arsonist's device?

   [ideation: that/ attitude: invoked positive VALUATION]
   Finesse: Embellish ⬇️ ?
   SUSPICIOUS FIRE BOND

2. So let me know what you think, **I want to hear** what <u>people's opinions</u> are on it

   [ideation: people's opinions/ attitude: positive AFFECT: DESIRE]
   Finesse: Embellish ⬇️ you think, I want
   Convoke: Marshal ⬇️ you
   LISTENING YOUTUBER BOND

In the first example, the use of question marks, makes the statement ambiguous. The YouTuber has a range of options to choose from for what supposedly caused the fire. In this sense, the audience is given more agency to decide rather than being told what is exactly in the video shown. There is no inscribed evaluation in this example, but there is invoked evaluation in the sense that by questioning what is occurring, the YouTuber expresses fascination (POSITIVE VALUATION). In the second example, the YouTuber asks

the audience what their opinions are on the video. By asking '*what you think*' and stating a '*I want to hear*' the YouTuber again opens up their statement to evaluation from outsiders. In this example we also see how MARSHALLING (the vocative you) and EMBELLISHING (the indefinite phrasing of think) work together in order to address an audience and leave a video open to interpretation.

These strategies of DISTILLING and EMBELLISHING were important across the dataset in terms of how social bonds were heteroglossically arranged. DISTILLING meant that social bonds were definitive without any room for negotiation, whilst EMBELLISHING meant that social bonds were negotiable and welcomed audience feedback. Due to the amount of hate speech contained in these videos, DISTILLING was a more common strategy as little room was given for changing ideologies. However, with videos that featured less ideology and more conspiracy (e.g. playing video clips of suspicious activity) EMBELLISHING was more frequently used, as the YouTuber would invite audiences to engage in conspiratorial thinking with them.

When YouTubers were communicating with their audience, values were upscaled or downscaled, as a means of explaining their importance or lack of importance. TEMPERING affiliation encompasses these ideas, in particular, how certain values attract attention. MODULATING was the more common strategy, emphasising the significance of claims:

3. This is a **really taxing** subject and it **made me angry** just writing this out.
   [ideation: subject/ attitude: negative APPRECIATION]
   [ideation: it/ attitude: negative REACTION]
   Temper: Modulate ↘ very & really
   EMOTIONAL EVENT BOND

4. This is just **crazy**, they **think we are really just mentally ill people**
   [ideation: this (event)/ attitude: negative REACTION]
   SUSPICIOUS FIRE BOND
   [ideation: they (media)/ attitude: negative VERACITY]
   LYING MEDIA BOND

[ideation: we/ attitude: negative CAPACITY]

INCAPABLE PEOPLE BOND

Temper: Modulate ↘ just & really just

Convoke: Marshal ↘ we

In these examples, MODULATING emphasises the key social bonds that are shared, whether that is emphasising the emotional nature of the subject (as in example 3 with the 'emotional event bond') or emphasising the suspicious nature of the fire (the negative evaluation of the event forming a 'suspicious fire bond'). Additionally, in example 4, the MODULATING draws further attention to the conspiracy the YouTuber is trying to propagate. The claim *'they think we are really just mentally ill people'* negatively evaluates the media (*they*) as spreading lies to what they believe is a naïve audience, that is evaluated with negative CAPACITY. The MODULATING has an important role here in extending this claim, with the repetition of 'just' extending the impact of the words the YouTuber is speaking.

In terms of genre structure, the social media voice-over macro-genre had particular genre staging. Firstly, the videos contained the channel logo or a social media post as the opening visual content. Secondly, this social media evidence became the main focus of the video, either live stream footage or screenshots being shown. This footage was zoomed in and out of by the YouTuber or repeatedly looped, in order to enhance their views with the audience. Online news articles were also spliced into the video in order to enhance the YouTuber's conspiratorial views or denounce mainstream media reporting. Lastly, the outro of the video featured the channel logo again and asked for subscribers. A diagram summarising these genre stages can be seen in Figure 6-18.

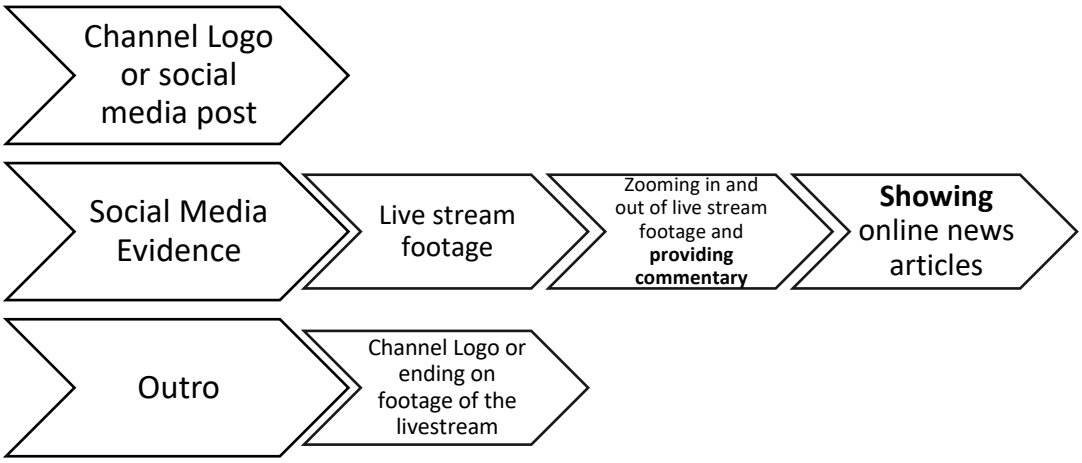


Figure 6-18 - Social Media Voice-Over Macro-Genre Structure

Visually, the social media voice-over macro-genre primarily contained technological authorisation as the key legitimation strategy or in some videos, as the only legitimation strategy. This was due to the absence of the YouTuber visually, as live stream footage comprised the entire frame. For example, in Figure 6-19 the video begins with a tweet that contained an embedded video. Throughout this video, the focus is on the YouTuber's computer screen. The YouTuber's voice is heard, but they are absent visually. Rather, the viewer follows the YouTuber in zooming into the embedded video. The YouTuber then points out the figure in the video and pauses and rewinds the footage multiple times. The video player bar is also prominently shown in the video, again, allowing the viewer to go on the journey with the YouTuber in terms of watching this footage and striking suspicion. In one frame, the footage quickly turns to a news article, with the computer screen again the focus. Overall, in this particular macro-genre live TV or looped video footage is the central focus point.

Figure 6-19 – Key Visual Legitimation Strategies in the Social Media Voice-Over Macro-Genre

Overall, the social media voice-over macro genre, was a common strategy in the dataset that predominately featured TECHNOLOGICAL AUTHORISATION as the key legitimation strategy in order to legitimate the bond that the Notre Dame fire is suspicious. This particular macro-genre was not complex in its genre-staging, in terms of not featuring many unique stages. Rather these videos were not particularly rehearsed and instead the YouTuber would encourage their audience to speculate with them in real-time.

6.4.2. Conspiracy Vlog Macro-Genre

In contrast, the conspiracy vlog was a more structured macro-genre that used particular rhetorical strategies to target their audience. Understanding how YouTuber's address their community provides insight into the type of the community the YouTuber is aiming to speak to, and how they want others to react to the key values they promote. In the Notre Dame Fire dataset, the majority of videos (14 out of 15) were directly

addressing a community in order to make their videos relatable and persuade their addressed community to align with their values. MARSHALLING was the most common CONVOKING affiliation strategy used, where vocatives such as *'you', 'folks' and 'y'all'* and forms of references such as *'listen here', 'you know'* and *'you can check'* were used to address or direct communication towards a community. For example, refer to the following short excerpts from across the transcripts:

5. See Brute **was live at Notre Paris** you can check all this for yourself
   [ideation: Brute/ attitude: invoked positive VALUATION]
   Convoke: Marshal ↘ you can check
   TRUTHFUL INFORMATION BOND

6. It is **pretty unusual** stuff so I'll play the video and you can see it
   [ideation: it (fire)/ attitude: positive VALUATION]
   Finesse: Distil ↘ it is
   Convoke: Marshal ↘ you can see
   SUSPICIOUS FIRE BOND

In all these examples, MARSHALLING is used to persuade the addressed community to align around the YouTuber's values. The proposition that '*you can check all this for yourself*' is used to support the YouTuber's assertion that the information they are providing is truthful because it can be verified. Many YouTubers in the dataset attempted to persuade their audience to align around the 'suspicious fire bond' with MARSHALLING affiliation as in '*you can see it*', that presumes there is evidence for the claims made. These strategies make these videos particularly deceiving, despite the supposed evidence shown being out of context or fabricated (for example, visually in these videos, tweets would be shown out of context).

DISTILLING was a common strategy used in this macro-genre. Videos that predominately used DISTILLING made definitive statements in order to limit any doubt that their conspiracy theory was incorrect. For example, in this extract from a YouTube video promoting a globalist conspiracy theory in relation to the Notre Dame Fire, a story is told about the meaning behind each world war, constructing a 'powerful wars bond':

7. The plan of course, the plan is you have World War one to **destroy the old order of Europe, the monarchies, Bazaar the Kaiser**, then you have World War two to **bring about like the UN to bring about the Jewish state Palestinian state**, and, and then World War three that's the **war of Christianity against Islam.**
   [ideation: World Wars/ attitude: POSITIVE VALUATION]
   Finesse: Distil ⬇️ the plan is
   POWERFUL WARS BOND
   [ideation: Islam/ attitude: NEGATIVE PROPRIETY]
   Finesse: Distil ⬇️ that's
   EVIL ISLAM BOND

In this example, there is repeated reference to '*a plan*' and a retelling of world wars. This constructs a story, connecting events together. The reference to '*world wars*' has POSITIVE VALUATION because these wars have brought about outcomes, so in other words, they are powerful. There is a lack of inscribed evaluation in this example, but Islam is still evaluated as evil at the end, with the DISTILLING affiliation strengthening the claims throughout this text by providing no other alternatives. By using these strategies, the YouTuber constructs their moral legitimation for why the West is currently threatened.

In terms of genre stages, the Conspiracy vlog was more structured compared to the social media voice-over macro-genre. These videos would all begin with footage of the vlogger introducing themselves. Screenshots of social media sites were then shown as evidence. The YouTuber would either go back and forth with footage of themselves and the evidence, or create a split screen structure in order to show themselves and the footage simultaneously. In comparison to the more ad-hoc footage of the social media voice-over, the conspiracy vlog was much more edited, with particular tweets and social media posts also selected and highlighted throughout the videos in order to emphasise claims like an evil Muslims bond (e.g., by stating Muslims had insulted white European culture). Lastly, the channel logo was shown and the YouTuber would promote their brand and ask for more subscribers. These genre stages are summarised in Figure 6-20.

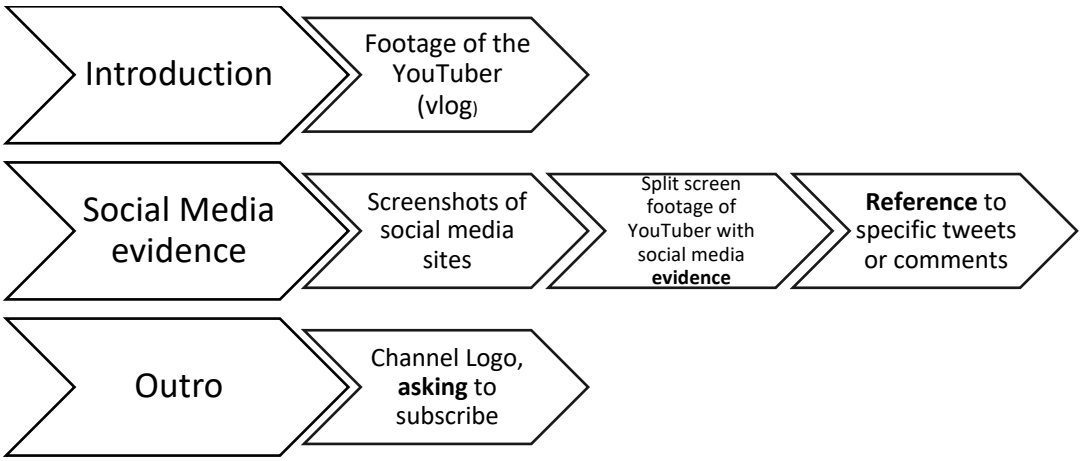


Figure 6-20 - Conspiracy Vlog Macro-Genre Staging

Visually, a range of legitimation strategies were evident in the conspiracy vlog. Mainly, this alternated between footage of the YouTuber (PERSONAL AUTHORISATION) and footage of the social media evidence (TECHNOLOGICAL AUTHORISATION). The social media evidence was typically zoomed in and out. Individually selected social media posts were also shown out of context, highlighting the very edited and curated nature of these videos. Lastly, branding was particular concern, with the large subscribe button at the end of the video representing MARKETING AUTHORISATION. These descriptions are summarised in *Figure 6-21*.

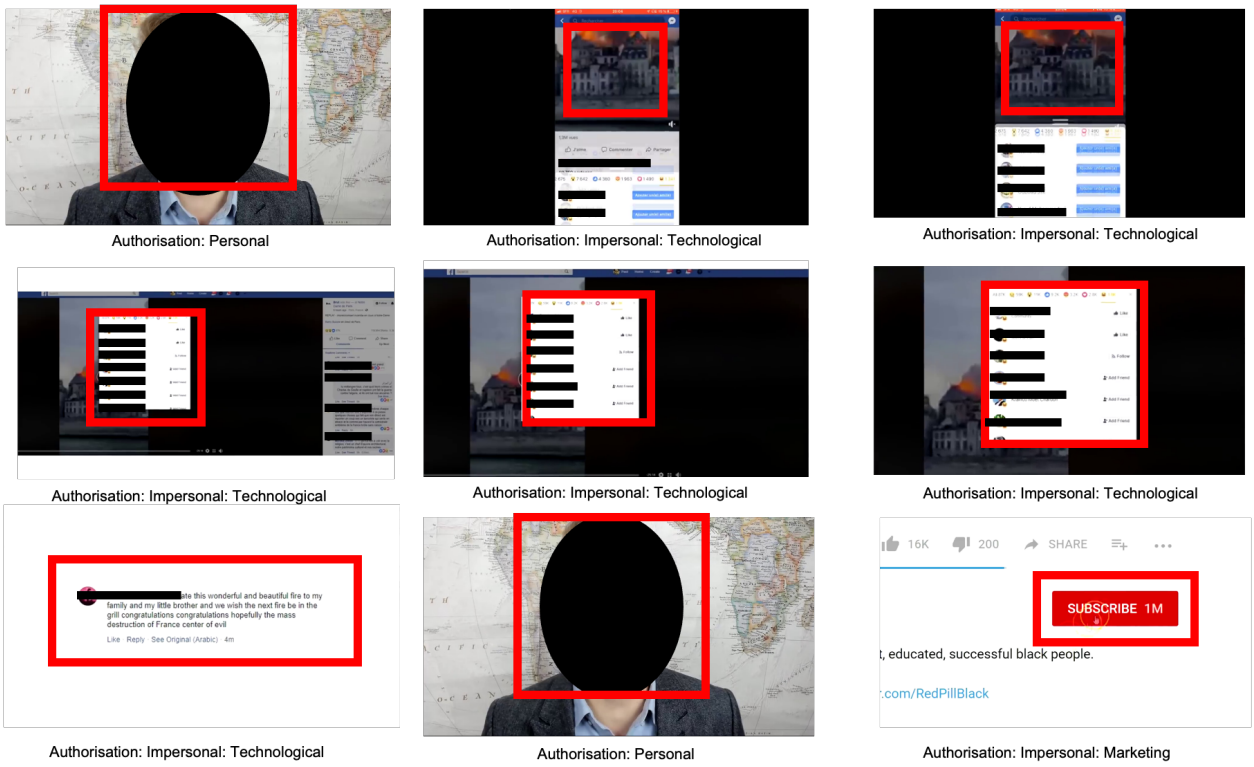


Figure 6-21 - Key Visual Legitimation Strategies in the Conspiracy Vlog Macro-Genre

Overall, the Conspiracy vlog macro-genre was more curated than the ad-hoc social media voice-over macro-genre. This macro-genre featured curated social media posts, and legitimation strategies that alternated between PERSONAL AUTHORISATION and TECHNOLOGICAL AUTHORISATION in order to elicit a suspicious fire bond. The social bonds in this macro-genre were also more targeted, going beyond the suspicious fire bond to also target minority groups in France such as Muslims.

6.5. Conclusion

This chapter has explored how social bonds in discourse are legitimated, both in language and through the visual modality. By using the Notre Dame Fire as a case study, this chapter has explored how YouTubers get people to align with their xenophobic and conspiratorial views. It has also shown how they legitimate their views through TECHNOLOGICAL AUTHORITY both visually and verbally, and shown via the analysis of the

transcripts, by legitimating MORAL EVALUATIONS, STORYTELLING and RATIONALISATION, and delegitimating AUTHORITIES and CONFORMITY.

The results of this study have shown how affiliation is an important strategy for persuading others to align with the sets of social values proposed in YouTube discourse. The communing affiliation analysis conducted on the video transcripts highlighted which affiliation strategies were more frequently used in order to position audiences to align with these values. These strategies include addressing a community, emphasising particular social bonds, and confining a social bond to a limited set of options.

The multimodal discourse analysis has demonstrated how YouTube videos can spread conspiratorial and hateful content in sophisticated ways, and how screenshots have contributed to the believability and virality of the particular videos discussed in this chapter. In other words, the assumption that screenshots solely act as evidence should not be taken for granted and the screenshot's context needs to be taken into account. This has several ethical implications for how we think about and use screenshots in our daily lives – have we taken for granted screenshots as evidence and ignored the broader social contexts in which screenshots are implicated?

7. CHAPTER SEVEN COMPARING CASE STUDIES: PERSONAE IN POLITICAL VERSUS NON-POLITICAL DISCOURSE

7.1. Introduction

This chapter is concerned with the construction of textual personae in the comment sections of videos about The Momo Challenge and The Notre Dame Fire. As discussed in *The Theoretical Framework Chapter*, the concept of a textual persona draws on Firth's notion of the persons and personalities that commune in discourse (Martin, 2009). Textual personae do not refer to individual people, but rather the generalisation of linguistically enacted identities. Firstly, the textual personae identified in the initiating comments of Momo Challenge videos will be explored via affiliation (Inwood and Zappavigna, 2021) and legitimation analyses. The replies to these comments are also analysed via a dialogic affiliation analysis. This is followed by an analysis of the textual personae identified in the Notre Dame Fire videos case study (Inwood and Zappavigna, 2023). The replies to these comments are also analysed via a dialogic affiliation analysis (Inwood and Zappavigna, 2022b). Further reflections are then made regarding the significance of textual personae and how the social bonds of textual personae can be visually represented.

After providing this overview into the personae detected in each case study, the non-political (Momo Challenge) and political (Notre Dame Fire) dimensions of textual personae will be discussed. This discussion will provide an opportunity to evaluate the usefulness of personae in understanding the various communities that spread discourses of misinformation and disinformation. Lastly, this chapter will discuss how the findings of this research are relevant across other domains and case studies.

The main research questions for the comment analysis were:

1. What are the main ideational targets and attitudes that users' bond around when discussing the Momo Challenge on YouTube? How are these social bonds legitimated?
2. By analysing the language of these comments, can a system of textual personae categories be developed? How does the notion of textual personae illuminate issues related to information disorders?
3. What patterns in evaluative language and affiliation strategies emerge when analysing replies to initiating comments?

7.2. Personae in the Momo Challenge Case Study

This section details the affiliation and legitimation analysis undertaken on 750 initiating comments and their replies sampled from the 15 Momo Challenge videos selected (refer to *The Methodology Chapter* for the sampling strategy). As stated in *The Methodology Chapter*, the analysis of main comments (here forth referred to as initiating comments) was firstly separated into comments that inferred 'Momo is real' versus comments that inferred 'Momo is fake' in order to obtain a dataset with a diverse range of comments. Within these two broad categories, the main values expressed through couplings of ideational and attitudinal meaning, and affiliation strategies were identified in order to understand the types of personae that reoccurred in the dataset. This is an example of how an initiating comment, and its replies are annotated (ideation underlined, attitude in bold and legitimation highlighted in grey):

Initiating comment:

I. do you have an example of a video that **actually has momo appear on it** what channel it came from?

SCEPTIC PERSONAE

[ideation: momo/ attitude: negative VERACITY]

Convoke: Marshal (do you) ↘️ SUPERIOR KNOWLEDGE BOND

Authorisation: Impersonal: Technological

Response comments:

i. No one does. It **doesn't exist**. It's a joke on **naive** adults.

[ideation: it (momo)/ attitude: negative VERACITY]

FAKE MOMO BOND ↘ MANAGE: SUPPORT: WARRANT: ADJUST

[ideation: adults/ attitude: negative CAPACITY]

MANAGE: SUPPORT: WARRANT: ADJUST ↘ IGNORANT ADULTS BOND

ii. Momo is a **social scare** that was originally a japanese art project that someone turned into a **worldwide global scare**

[ideation: momo/ attitude: negative VALUATION]

MANAGE: SUPPORT: WARRANT: RALLY ↘ FAKE MOMO BOND

[ideation: momo/ attitude: negative REACTION]

MANAGE: SUPPORT: WARRANT: RALLY ↘ SAFE MOMO BOND

Refer to *The Methodology Chapter* for a further explanation of the annotation strategy. For the replies, the focus was on dialogic affiliation, hence the different annotation style to the initiating comments.

7.2.1. Analysis of Initiating Comments: Communing Affiliation and Legitimation

The affiliation and legitimation analysis undertaken on the initiating comment corpus revealed six key textual personae who each manifest a particular patterning of couplings and tendency towards sharing or negotiating particular bonds. These personae embodied different values and ways of affiliating. This section will illustrate the findings from this analysis, and its contribution to understanding the values at stake when people respond to internet hoaxes.

7.2.1.1. Connoisseurs

Connoisseurs, that is, personae that characterised itself as knowledgeable about internet culture, rallied around social bonds that stated 'Momo is a statue' or 'Momo is

a meme'. These sorts of statements were mostly factually correct. The image of 'Momo' originated from a sculpture created by the Japanese artist Keisuke Aiso who worked for the Link Factory, a Japanese special effects company[22]. This image of the sculpture was later appropriated by the Momo Challenge internet hoax without reference to the creator of the original sculpture from which Momo's visual depiction was derived. Thus, many of the comments were alluding to this history. Connoisseurs would DISTIL the opinion that Momo is a benign object rather than a real-life figure that is a danger to society. In the comment sections, Connoisseurs used RATIONALISATION strategies to explain the meaning behind the origins of Momo:

1. For anyone this actually freaks out, "Momo" is actually **a cropped image of the statue "Mother Bird" made by the Japanese special effects company Link Factory**. It is not **going to hurt you** in any way. It was **made for an art exhibition called "Ghosts and Spectators"** hence the **creepy** nature of the statue.

   [ideation: Momo/ attitude: positive APPRECIATION]

   [ideation: Momo/ attitude: positive PROPRIETY]

   [ideation: Momo/ attitude: positive APPRECIATION]

   [ideation: statue/ attitude: negative REACTION]

   Finesse: Distil (actually) ↘ SAFE MOMO BOND

   Convoke: Designate (for anyone, you) ↘ community who fear Momo (DANGEROUS MOMO BOND)

   Temper: Foster (in any way, hence) ↘ SAFE MOMO BOND

   Rationalisation: Theoretical: Explanation

This comment uses a CONVOKING affiliation strategy ("*for anyone this actually freaks out*") to directly address people who are '*freaked out*' about Momo, therefore outgrouping non-Connoisseurs. The user acknowledges that the statue is '*creepy*' (negative REACTION), with the key distinction being that momo is considered a '*statue*' not a person. There are two competing bonds in this comment: a SAFE MOMO BOND that assures the reader that Momo is just a statue and a DANGEROUS MOMO BOND held by those 'freaking out' about Momo. In terms of legitimation strategies, RATIONALISATION:

[22] The fact-checking site Snopes provides information about momo's origins https://www.snopes.com/news/2019/02/26/momo-challenge-suicide-game/

THEORETICAL: EXPLANATION is predominately used as the commenter explains that momo is '*a cropped image*' and was '*made for an art exhibition*', therefore explaining its creepy nature. Thus, Connoisseurs promoted research and rational thought in their responses.

Another pattern of negative REACTION towards the Momo statue and a FAKE MOMO BOND was common across many comments such as the following:

2. Momo is a statue. It's a piece of art. Art can be **creepy**.
   [ideation: art/ attitude: negative REACTION]
   Finesse: Embellish (can be) ↘ FAKE MOMO BOND
   Rationalisation: Theoretical: Explanation
   Moral Evaluation: Evaluation

With this comment, the 'creepiness' of the artwork is stressed in terms of negative REACTION, again moving away from treating Momo as a person, to instead treating Momo as, a statue. '*Art can be creepy*' is not a definitive statement, thus the '*can be*' is an expansive statement, entertaining a thought about art's creepiness, thus EMBELLISHING the FAKE MOMO BOND. These sorts of comments acknowledged the thoughts many people had about Momo's creepiness but tried to counteract false information about the Momo challenge by asserting that Momo is an object, not a person that is a danger to society. Again, RATIONALISATION: THEORETICAL: EXPLANATION is a common strategy in explaining that momo '*is*' a statue. In this example, MORAL EVALUATION: EVALUATION is also used as a strategy to ascertain that it is normal for some art to appear *'creepy'*.

Another focus of Connoisseurs was highlighting the negative VERACITY of the videos they were commenting on. For instance, this comment directly targets the video by using negative judgement, and a mix of TEMPERING and DISTILLING affiliation resources:

3. Videos such as this one just **perpetuate this hoax**...do your homework...this has been proven **fake**, and not one death has been attributed to it. "Momo"; is **a statue** created by a Japanese artist.

    [ideation: video/ attitude: negative VERACITY]

    [ideation: momo/ attitude: negative VERACITY]

    [ideation: "Momo"/ attitude: positive APPRECIATION]

    Finesse: Distil (not one death, is) ↘️ SUPERIOR KNOWLEDGE BOND & SAFE MOMO BOND

    Temper: Modulate (just) ↘️ FAKE MOMO BOND

    Convoke: Marshal (your) ↘️ SUPERIOR KNOWLEDGE BOND

    De-Authorisation: Impersonal: Technological

    Rationalisation: Theoretical: Explanation

There is negative JUDGEMENT towards the video creator. '*Just*' acts as a form of TEMPERING to intensify the claim that the Momo Challenge is a hoax. By referring to '*Videos such as this one*' TECHNOLOGICAL AUTHORITY is relied upon. The FINESSING affiliation strategy expands the coupling by stating that not one death has been attributed to the Momo Challenge, creating a SUPERIOR KNOWLEDGE BOND as the commenter has done their '*homework*' and a SAFE MOMO BOND. Again, RATIONALISATION: THEORETICAL: EXPLANATION is used as a strategy to explain that Momo is fake. Similar to the previous comments, Momo also switches from being considered as a person to just being a '*statue*'.

Another key concern for Connoisseurs was highlighting the Momo Challenge's origins as a meme and creepypasta. A creepypasta refers to a type of fictional horror story that is shared on the internet, on sites such as Reddit. Thus, when Momo is described as a '*meme*' or '*creepypasta*' by Connoisseurs, they are judging Momo in terms of invoked negative VERACITY, as explored in these comments:

4. The momo thing is **fake**. If you go onto know your meme, it explains how it's **a sculpture on a stand'n shit.**

   [ideation: momo/ attitude: negative VALUATION]

   [ideation: it (momo)/ attitude: invoked positive APPRECIATION]

   Finesse: Embellish (it explains) ↘ FAKE MOMO BOND

   Temper: Modulate ('n shit) ↘ FAKE MOMO BOND

   Convoke: Marshal (you) ↘ SUPERIOR KNOWLEDGE BOND

   Authorisation: Impersonal: Technological

5. Wait why is Momo becoming a **big internet creepypasta thing** now when the **meme** is like 7 months old

   [ideation: momo/ attitude: negative T-VERACITY]

   Temper: Modulate (wait) ↘ FAKE MOMO BOND

   Finesse: Embellish (why) ↘ SUPERIOR KNOWLEDGE BOND

   De-Authorisation: Impersonal: Technological

In these comments, Momo's origins are stated as a '*meme*'. '*Know your meme*' is a popular site detailing the origins of popular memes and is referenced in the first comment, EMBELLISHING a FAKE MOMO BOND and acting as a form of TECHNOLOGICAL AUTHORITY. By explaining that '*If you go onto know your meme*' CONVOKING affiliation is also used, establishing a SUPERIOR KNOWLEDGE BOND due to the commenter's knowledge of internet culture. In the second comment, Momo as a "*meme*" is portrayed as insignificant in comparison to Momo as a creepypasta which has become a '*big internet…thing*'. Through the use of '*why*' EMBELLISHING affiliation is used to expand the coupling and form a SUPERIOR KNOWLEDGE BOND as the commenter questions the sudden resurgence of the fake Momo Challenge. In this instance, TECHNOLOGICAL AUTHORITY is used as a delegitimation strategy, as the focus on Momo as a meme elicits its negative VERACITY.

Ultimately, Connoisseurs are distinguished by their persuasive argument that Momo is an object not a person and their knowledge of internet terminology such as creepypastas and their fictitious origins. Additionally, Connoisseurs clashed with the shared values in the YouTube videos regarding whose interpretation of Momo was more legitimate.

This latter position involved negotiating bonds about how true or false, or safe or dangerous, Momo was deemed to be. RATIONALISATION and TECHNOLOGICAL AUTHORITY were the most common legitimation strategies to provide evidence for these claims.

7.2.1.2. Nationalists

The Nationalist persona similarly centred fake Momo bonds but also stressed the importance of nationality. Broadly, a Nationalist refers to someone who is devoted to their nation (Harper, n.d.). But a Nationalist can also be considered as a term for someone who belongs to a particular nationality or shows support for national interests (Harper, n.d.). CONVOKING affiliation was used by Nationalists, in order to name a particular country and get people to bond around the attitudes expressed by that country. Nationalists were eager to present their country as more technologically knowledgeable than other countries. This can be realised in the following comments:

6. Here in Philippines momo is just a **meme** for us
   [ideation: momo/ attitude: invoked positive AFFECT]
   Convoke: Designate (Philippines, us) ↘ SUPERIOR NATION BOND
   Temper: Modulate (just) ↘ SAFE MOMO BOND
   De-Authorisation: Impersonal: Technological
   Authorisation: Personal

7. It's **funny** to hear this because here in Brazil momo became a **meme**.
   [ideation: this (proposition)/ attitude: negative REACTION]
   [ideation: momo/ attitude: invoked positive AFFECT]
   Convoke: Designate (Brazil) ↘ SUPERIOR NATION BOND
   Finesse: Distil (because) ↘ FAKE MOMO BOND
   De-Authorisation: Impersonal: Technological
   Authorisation: Personal

In these comments, CONVOKING affiliation is used to name certain communities – '*here in the Philippines*' and '*here in Brazil*'. Momo is downplayed as being '*just a meme*' – a form of negative VALUATION if momo is considered as an object, or negative CAPACITY, in the instance where momo is spoken of as if she is a person. In terms of legitimation

strategies, these sorts of comments legitimate the PERSONAL AUTHORITY of the people who represent the communities named, and delegitimate the TECHNOLOGICAL AUTHORITY of the meme. Rather than being scared of Momo as US-centric YouTube videos often portrayed, these users are inferring that Momo is just a joke to them. This forms a SUPERIOR KNOWLEDGE BOND where through different affiliation strategies certain countries can out-group US interpretations of Momo.

Whilst these comments were relatively short, some longer comments would add multiple targets of evaluation, but followed the same CONVOKING affiliation strategy. Take for instance, this comment:

8. **Haha** 😂😂😂 in Latin America, last year this "trend" went viral and now no one is **scared** by it. I am surprised that momo arrived in America too! **L** and momo supposedly speaks Spanish. And it is a Japanese number... and now it turns out that it speaks English? This is **crazy stuff L**

    [ideation: user's attitudes to momo/ attitude: positive SECURITY]

    [ideation: momo spreading to other countries/ attitude: negative DISSATISFACTION]

    Convoke: Designate (Latin America) ↘️ SUPERIOR NATION BOND

    Finesse: Embellish (no one, supposedly) ↘️ FAKE MOMO BOND

    Authorisation: Personal

    Authorisation: Impersonal: Technological

    Rationalisation: Theoretical: Explanation

'*In Latin America'* CONVOKES a community that is no longer scared by Momo, similar to the previous comments in their establishment of a SUPERIOR NATION BOND. However, an extra reflection is added by expressing DISSATISFACTION at Momo spreading to other countries. FINESSING affiliation is used to expand the second coupling that Momo '*supposedly*' speaks Spanish but is also intertwined within Japanese and English contexts. This invokes a FAKE MOMO BOND due to the geographical inconsistencies associated with Momo. These sorts of longer comments feature multiple legitimation strategies; the PERSONAL AUTHORISATION of the people representing the countries named, the TECHNOLOGICAL AUTHORITY references regarding internet culture, and the

RATIONALISATION strategies used to explain the illogical nature of the claims being made about the Momo Challenge.

Nationalists also prallied around a dichotomy between Americans as '*anxious*' and non-Americans as informed and capable in handling Momo. For instance, this comment is more generalised but the '*some countries*' could be insinuated as "*Americans*" (DISINCLINATION) whilst '*in the philippines*' people are not scared of momo and therefore are evaluated as having positive SECURITY:

9. Why america is **afraid of momo**? Here in the philippines were **not afraid of momo** we just **created meme for her**

    [ideation: America/ attitude: negative DISINCLINATION (FEAR)]

    [ideation: we in the Philippines/ attitude: positive SECURITY (CONFIDENCE)]

    [ideation: we in Philippines/ attitude: invoked positive ATTITUDE]

    Convoke: Marshal (here in the Philippines, we) ↘ SUPERIOR KNOWLEDGE BOND

    Finesse: Embellish (why & ?) ↘ SAFE MOMO BOND

    Temper: Modulate (just) ↘ SAFE MOMO BOND

    Authorisation: Personal

    Authorisation: Impersonal: Technological

This comment therefore has a safe Momo bond by the question '*why America is afraid*' expanding the coupling. A SUPERIOR NATION BOND also exists by MARSHALLING a community of Nationalists who positively associate Momo with '*safe*' memes. Again, the PERSONAL AUTHORITY of the country named '*philippines*' is used as a legitimation strategy, in conjunction with the TECHNOLOGICAL AUTHORITY expressed by understanding that Momo is just a meme.

Overall, the comments made by Nationalists followed a similar trajectory where commenters bonded around SUPERIOR NATION BONDS through CONVOKING affiliation strategies that named their own country. Additionally, evaluative language was used to comment on the harmlessness of the Momo Challenge, or to criticise Americans for taking the Momo Challenge seriously. In the YouTube comment corpus, despite the ability to remain anonymous regarding national identity, YouTubers frequently

discussed attitudes their country supposedly held towards Momo with PERSONAL AUTHORITY and TECHNOLOGICAL AUTHORITY legitimation strategies and criticised US-centric YouTube videos that promoted anxiety about the Momo Challenge.

7.2.1.3. Sceptics

YouTube comments that inferred 'Momo is fake' tended to use questions as a way of expressing this belief and expressed SUPERIOR KNOWLEDGE, SAFE MOMO, or FAKE MOMO bonds. Etymologically, the term 'sceptic' originated from an Ancient Greek school of thought that "doubted the possibility of real knowledge" (Harper, n.d.). Regarding textual personae, 'sceptic' is instead considered in terms of its more contemporary meaning, that is, one who "investigates or researches" as opposed to one who "asserts and thinks that he has found" (de Unamuno, 1924; Harper, n.d.). Sceptic comments either negatively evaluated Momo or positively evaluated the action of searching for evidence.

10. I would **love** to **investigate** this, does anyone know where I can find Momo's number or official videos? Have they all been taken down?
    [ideation: act of investigating/ attitude: positive desire]
    [ideation: this (Momo's number)/ attitude: invoked negative VALUATION]
    Convoke: Marshal (anyone) ↘ SUPERIOR KNOWLEDGE BOND
    Finesse: Embellish (would, ?) ↘ SAFE MOMO BOND
    Authorisation: Impersonal: Technological

11. I would **like** the link to the **actual** peppa pig/momo video.
    [ideation: access to link/ attitude: positive desire]
    [ideation: peppa pig/momo video/ attitude: negative VERACITY]
    Finesse: Embellish (I would) ↘ SUPERIOR KNOWLEDGE BOND & FAKE MOMO BOND
    Authorisation: Impersonal: Technological

For example, in the first comment the user invokes negative VALUATION towards the Momo Challenge by suggesting that this challenge should be investigated. Although they do not explicitly state Momo is fake, the act of questioning Momo goes against the

content in the YouTube video that explicitly states Momo is real. Therefore, a SUPERIOR KNOWLEDGE and SAFE MOMO BOND are shared in this instance. In the second comment, Momo is more explicitly inferred to as fake, as the user states they want to see an '*actual*' video, thus critiquing the supposed peppa pig/momo clip shown in the YouTube video being commented on. The user's desire for evidence and critique of the YouTube video creates a SUPERIOR KNOWLEDGE BOND and FAKE MOMO BOND. In these sorts of comments, TECHNOLOGICAL AUTHORITY (asking for Momo's number or video links) is relied upon as the only method for discovering whether Momo is real. This type of questioning that asks for an '*actual*' video as a TECHNOLOGICAL AUTHORITY legitimation strategy is also shown in the following comments:

12. do you have an example of a video that **actually** has momo appear on it what channel it came from?

    [ideation: momo/ attitude: negative VERACITY]

    Convoke: Marshal (do you) ↘ SUPERIOR KNOWLEDGE BOND & FAKE MOMO BOND

    Authorisation: Impersonal: Technological

13. You'd think there would be RECORDINGS of momo accessible online. Where are they? Where is the evidence that <u>this</u> is actually a **thing**?

    [ideation: (no recordings)/ attitude: invoked negative valuation]

    [ideation: this (momo)/ attitude: negative valuation]

    Convoke: Marshal (you) ↘ SUPERIOR KNOWLEDGE BOND

    Finesse: Embellish (you'd think) ↘ SUPERIOR KNOWLEDGE BOND

    Temper: Foster (RECORDINGS) ↘ FAKE MOMO BOND

    Authorisation: Impersonal: Technological

Both comments question whether the videos actually contain a recording of Momo, thus inferring that Momo is fake primarily through CONVOKING affiliation. In the first comment, the YouTuber is directly asked if they have an example of a video. In the second comment, heteroglossic language is used as in '*you'd think*' and '*where is the evidence*' EMBELLISHING a SUPERIOR KNOWLEDGE BOND. 'RECORDINGS' also acts as an example of TEMPERING – intensifying the demand for evidence and creating a FAKE MOMO

BOND. Sceptics also used TEMPERING affiliation in order strengthen the evaluation of Momo as fake, as in the following example:

14. So why does this one video "proof" not show the video that the momo clip was supposedly cut into? This can be so easily **faked.**

    [ideation: momo clip/ attitude: negative *t*-VERACITY]

    Finesse: Embellish (supposedly) ↘ SUPERIOR KNOWLEDGE BOND

    Temper: Modulate (so easily) ↘ FAKE MOMO BOND

    De-Authorisation: Impersonal: Technological

Again, the heteroglossic voice construes the meaning that the Momo clip '*was supposedly cut into*' forming a SUPERIOR KNOWLEDGE BOND and delegitimating the TECHNOLOGICAL AUTHORITY of the supposed video proof. TEMPERING also embodies the coupling of momo as negative VERACITY, in that momo is '*so easily*' faked, thus creating a FAKE MOMO BOND. In all, what can be derived from the language of Sceptics is a focus on asking questions that would hold negative, invoked evaluations of momo as fake and how Sceptics position themselves as a select group of people holding a SUPERIOR KNOWLEDGE BOND with the ability to primarily use TECHNOLOGICAL AUTHORITY legitimation strategies to justify their claims.

7.2.1.4. Critics

The last main personae that held a 'Momo is fake' bond, were Critics, with these critics insinuating that the media and older generations misunderstand internet culture, and instead spread unnecessary hysteria and moral panic. In its original definition, the Critic, as a noun, refers to a person who analyses, evaluates or comments on literary texts (Harper, n.d.). Overtime, a Critic has been broadened to refer to more generally someone who expresses considered judgement, especially that which is harsh or unfavourable (Harper, n.d.). It is this second definition that the Critic persona in the Momo Challenge case study most aligns with. In many of the comments by the Critic persona "*moral panic*" was used as an attitudinal target:

15. It is the **most talked about viral scare story** of the year so far, blamed for child suicides and **violent attacks** – but experts and charities have warned that the "Momo challenge" is nothing but a "**moral panic" spread by** adults.

    [ideation: it is (Momo challenge)/ attitude: positive APRECIATION]

    [ideation: it is (Momo challenge)/ attitude: negative *t*-VERACITY]

    [ideation: it is (Momo challenge)/ attitude: negative *t*-PROPRIETY]

    [ideation: adults/ attitude: negative CAPACITY]

    Temper: Modulate (most) ↘️ FAKE MOMO BOND

    Finesse: Embellish (experts and charities have warned, nothing but) ↘️ BAD ADULT BOND

    Authorisation: Expert

    Rationalisation: Theoretical: Explanation

16. **Warning**⚠️The news media just discovered Momo. Mass hysteria **overreaction** to **misunderstood internet culture** imminent. Prepare for **wide-spread moral panic, demonetization** video takedowns.

    [ideation: (story)/ attitude: invoked positive VALUATION]

    [ideation: news media/ attitude: invoked negative CAPACITY]

    [ideation: (media)/ attitude: negative CAPACITY]

    [ideation: moral panic / attitude: negative VALUATION]

    Temper: Modulate (⚠️, just) ↘️ FAKE MEDIA BOND

    Moral Evaluation: Evaluation

    De-Authorisation: Impersonal: Technological

In the first comment, by describing the Momo Challenge as a '*viral scare story*' or '*moral panic*', the Momo Challenge is judged with negative VERACITY and a FAKE MOMO BOND with the tempering affiliation of '*most talked about*'. '*Experts and charities have warned*' represents an EXPERT legitimation strategy and uses a FINESSING affiliation strategy, acknowledging other voices in discourse and forming a BAD ADULT BOND. RATIONALISATION is also used as a legitimation strategy to ascertain the illogical nature of the claims made about the Momo Challenge. In the second comment, moral panic can be associated with negative VALUATION and MORAL EVALUATION as a legitimation strategy. '*Mass hysteria*' is associated with the negative CAPACITY of internet users and the de-legitimation of the news media. The TEMPERING affiliation of '*just*' establishes a predominant FAKE MEDIA BOND.

In other comments, such as the following, America becomes the target of evaluation, associated with '*hysteria*':

17. **Zero deaths** have actually proven to be linked to blue whale or momo. No one actually knows if the Argentina girl was linked to momo. But of course the media **sensationalizes** these **dumb** challenges because it makes **great** headlines and **scary** pictures. America **hysteria**

    [ideation: blue whale or momo/ attitude: negative VERACITY]

    [ideation: media/ attitude: negative VERACITY]

    [ideation: america/ attitude: negative VERACITY]

    Temper: Modulate (of course) ↘ BAD MEDIA BOND

    Finessing: Embellish (no one actually knows) ↘ FAKE MOMO BOND

    Rationalisation: Theoretical: Explanation

    De-Authorisation: Impersonal: Technological

    Moral Evaluation: Evaluation

Americans and American news media are evaluated with negative VERACITY and a mix of TECHNOLOGICAL DELEGITIMATION and MORAL EVALUATION strategies. '*Actually proven*' and 'of course' are used as TEMPERING affiliation and RATIONALISATION strategies, strengthening the claims of a FAKE MOMO BOND and BAD MEDIA BOND. In addition, the Momo Challenge and blue whale challenge are both associated with negative VERACITY, grouped together as internet hoaxes that have both been sensationalised by media outlets that don't understand internet culture.

The final main type of comment that Sceptics make, involves criticising the older generation and providing a direct comparison to the younger generations:

18. God the older generation is **stupid**…adults **believe nearly anything without proof**, I'm glad younger people are actually starting to become at least a little bit **smart**. The momo challenge is **fake**, it doesn't exist. **No child has ever been contacted** by 'momo' or hurt themselves/killed themselves because of it. It was just **made up** by someone looking for clicks and now it's become way **too big of a thing**. It's not even **real**. You would think schools and news channels **would do a little research** **before spreading false information like crazy.**

    [ideation: older generation/ attitude: negative CAPACITY]

    [ideation: adult/ attitude: negative CAPACITY]

    [ideation: younger people/ attitude: positive CAPACITY]

    [ideation: Momo Challenge/ attitude: negative VALUATION]

    [ideation: Momo/ attitude: invoked positive JUDGEMENT]

    [ideation: it (Momo)/ attitude: negative VERACITY]

    [ideation: it (Momo)/ attitude: negative VALUATION]

    [ideation: it (Momo)/ attitude: negative VALUATION]

    [ideation: schools and news channels/ attitude: negative CAPACITY]

    Temper: Modulate (God, nearly anything, actually, at least a little bit) ↘️ BAD OLDER PEOPLE/ GOOD YOUNGER PEOPLE BOND

    Finesse: Distil (actually) ↘️ BAD OLDER PEOPLE/ GOOD YOUNGER PEOPLE BOND

    Temper: Modulate (has ever, just, too big, not even) ↘️ FAKE MOMO BOND; SAFE MOMO BOND

    Temper: Modulate (little) ↘️ BAD MEDIA/ SCHOOLS BOND

    De-Authorisation: Role Model

    Rationalisation: Theoretical: Explanation

In this comment, the older generation is associated with negative CAPACITY, due to their lack of media literacy skills (‘*believe nearly anything without proof*’). In comparison, young people are treated with positive CAPACITY (‘*little bit smart*’). Schools and news channels are associated with negative VERACITY in this comment, for being unable to do a ‘*little research*’ and ‘*spreading false information like crazy*’. Thus, this comment FOSTERS an opposition between a BAD OLDER PEOPLE and a GOOD YOUNGER PEOPLE BOND. Older generations and schools are delegitimated as role models, and the explanation of internet culture ‘*someone looking for clicks*’ serves as a RATIONALISATION strategy for legitimating the commenter’s claims. Additionally, FAKE MOMO BONDS and SAFE MOMO BONDS are shared as a critique of the media and schools. Overall, Critics would rely on a range of negative judgemental language to highlight how older generations and the

media fail to understand internet culture and spread unnecessary moral panic, due to their inability to properly research internet cultures.

7.1.2.5. Myth Spreaders

Myth Spreaders were a common persona in the 'Momo is real' grouping. These commenters focused on spreading stories about suicides linked to the Momo Challenge. A myth refers to an untold story or rumour that is widespread (Harper, n.d.). Etymologically, a myth spreader delivers their message "by word of mouth" (Harper, n.d.). Traditionally, myths were associated with supernatural beings and provided a justification for either the early history of society, a religious belief, or natural phenomenon (Harper, n.d.). In the Momo Challenge dataset, the Myth Spreader persona repeats messages that they have overheard of people supposedly killing themselves over the Momo Challenge. Many of these comments were more monoglossic in tone compared to other comments, in the sense that they tried to avoid directly naming other voices and presented their comments without entertaining other possibilities, as in the following examples:

19. UPDATE: A 12 year old girl in Argentina had committed suicide after following "**instructions**" from Momo. Kind of like Blue Whale.
    [ideation: Momo/ attitude: invoked negative PROPRIETY]
    Temper: Modulate (UPDATE) ↘ REAL MOMO BOND
    Finesse: Distil (Quotation marks, kind of) ↘ DANGEROUS MOMO BOND
    Mythopoesis: Cautionary Tale

20. It's **real** in the philipines 11 years old kid die because of the momo suicide challenge and blue wale challenge
    [ideation: momo/ attitude: positive VERACITY]
    Convoke: Designate (in the Philipines) ↘ REAL MOMO BOND
    Mythopoesis: Cautionary Tale

In the first comment, TEMPERING affiliation (UPDATE) is used to attract attention to the coupling of momo and negative PROPRIETY, creating a REAL MOMO BOND. There is no

explicit evaluation in the comment, only invoked evaluation as "MOMO" and "BLUE WHALE" are known as suicide games, therefore invoking negative PROPRIETY and a DANGEROUS MOMO BOND. In the second comment, '*it's real*' is the only form of evaluative language used, treating momo with positive VERACITY. '*In the philipines*' can be seen as a CONVOKING affiliation strategy, in terms of naming a community and adding to a REAL MOMO BOND. In all, these two comments are not as rich in evaluative language and affiliation strategies as other comments in the dataset, a key indicator of the sorts of comments written by Myth Spreaders. Additionally, these comments are characterised by MYTHOPOESIS: CAUTIONARY TALE as a legitimation strategy, in other words, using hearsay or narrative strategies to warn others about the Momo Challenge.

In comments by Myth Spreaders that did use more heteroglossic language, these comments all used Momo as the central ideational target:

21. I heard that this 4 year old commit suicide cuz Momo told him to do something really **dangerous**

    [ideation: impact of momo/ attitude: negative PROPRIETY]

    Temper: Foster (really) ↘ DANGEROUS MOMO BOND

    Mythopoesis: Cautionary Tale

22. I'm here because **a 10 year old boy died on Namibia this week because of** the momo challenge

    [ideation: momo/ attitude: negative JUDGEMENT]

    Finesse: Distil (I'm here because, because of) ↘ DANGEROUS MOMO BOND

    Mythopoesis: Cautionary Tale

In the first comment, heteroglossic language is used as the commenter writes that '*I heard*', attributing some voice to the discourse. TEMPERING affiliation is used to strengthen the claims of Momo's negative PROPRIETY in supposedly leading a child to suicide. In the second comment, similar to that of the Nationalist persona, a CONVOKING affiliation strategy is used '*I'm here because*', trying to align the comment with a community that cares about reporting suicides in relation to the Momo Challenge. Again, these comments solely adopted a MYTHOPOESIS: CAUTIONARY TALE legitimation

strategy, as stories with warnings are relied upon. Nonetheless, whilst these comments would have legitimation and affiliation strategies, there was not much consistency in terms of location and age of the children who had supposedly committed suicide. Thus, within a mixed affiliation analysis and corpus linguistics approach, it can also be revealed how inconsistencies in information provided by comments, provides an indication of the falsity of many of these comments.

The last main category of comments that Myth Spreaders wrote involved discussing what was overheard at school:

23. Apparently the principle in my school caught two kids cutting themselves with pencils in the bathroom because of the Momo challenge, and sent an email warning all parents about Momo. This is some **scary** shit.

    [ideation: Momo challenge/ attitude: invoked negative JUDGEMENT]

    [ideation: this shit/ attitude: negative REACTION]

    Finesse: Embellish (apparently) ↘ DANGEROUS MOMO BOND

    Temper: Foster (warning all) ↘ DANGEROUS MOMO BOND

    Mythopoesis: Cautionary Tale

With this comment, '*Apparently*' as an EMBELLISHING strategy (due to its heteroglossia) signals that these users have overheard something and are about to share more information in the comment, again relying on MYTHOPOESIS as a legitimation strategy. Momo is the main ideational target in this comment, with negative REACTION being directed towards the Momo Challenge and a DANGEROUS MOMO BOND.

Overall, Myth Spreaders did not switch ideational targets, but rather focused all their evaluations on Momo. Compared to the other 'Momo is fake' personae, Myth Spreaders wrote shorter comments often with monoglossic or minimal voices in discourse, relied on MYTHOPOESIS (the art of storytelling) as a legitimation strategy, wrote inconsistent statements, and relied heavily on TEMPERING affiliation strategies to attract attention to their comments.

7.2.1.6. Moralisers

Moralisers were preoccupied with providing negative moral JUDGEMENT towards hackers using technology for evil reasons, and towards the behaviour of parents or platforms such as YouTube. To moralise means to 'expound or interpret spiritual or moral significance' or to 'draw a moral from' (Harper, n.d.). Thus, a moraliser is given to making moral judgements, that is judgements about the right behaviour that needs to be exhibited. In many of these comments, the focus was on evaluating hackers, the feelings of children, parents, or platforms, rather than evaluating the Momo Challenge itself.

24. Hackers **hide behind** the Momo challenge to **black mail** kids. It's more **worse** than these kids videos. The WhatsApp Momo challenge could be **hidding pedophiles** and **sadistic hackers**.

    [ideation: hackers/ attitude: negative PROPRIETY]

    [ideation: it (hacker's behaviour)/ attitude: negative PROPRIETY]

    [ideation: WhatsApp Momo Challenge/ attitude: negative PROPRIETY]

    Temper: Foster (more worse) ↘️ BAD HACKER BOND

    Finesse: Embellish (could be) ↘️ DANGEROUS MOMO BOND

    Mythopoesis: Moral Tale

    Moral Evaluation: Evaluation

In this comment, hackers and paedophiles are evaluated with negative PROPRIETY and the EMBELLISHING of '*could be*' forms a DANGEROUS MOMO BOND. By recounting what hackers are doing, MYTHOPOESIS is used to legitimate the moral significance of the issue at hand. Hackers are expressed as being '*more worse*' than the YouTube kids videos that have Momo spliced into them. '*More worse*' situates this comment with a TEMPERING affiliation strategy and a BAD HACKER BOND, legitimating the moral significance of the issue. TEMPERING affiliation is also used in the following comment:

25. That's **creepy** af. I do absolutely believe there is some **creepy** psycho on the dark web trying to kill kids.

    [ideation: that's (momo)/ attitude: negative REACTION]

    [ideation: psycho on the dark web/ attitude: negative PROPRIETY]

    Temper: Modulate (af, absolutely) ↘ DANGEROUS MOMO BOND & BAD HACKER BOND

    Mythopoesis: Moral tale

In this comment, people on the dark web are evaluated with negative PROPRIETY. '*Absolutely*' strengthens the coupling and thus is another example of TEMPERING affiliation and a DANGEROUS MOMO and BAD HACKER BOND. Again, by repeating hearsay that there is '*some creepy psycho on the dark web*' narrative strategies are used as MYTHOPOESIS: MORAL TALE legitimation.

Moralisers would also write comments that stated, 'momo is real' and would align parents with negative judgement. Often these comments would not centre evaluation on momo but would rather centre evaluation about the feelings of children and/or judgement about parents. For instance, in these comments, parents are treated with negative CAPACITY due to their laziness when it comes to looking after their children:

26. Oh no! now all u **lazy** parent can't use YouTube to entertain ur kids. Now it **gna have to get involved with ur kid life's Instead of leaving them in the other room on the iPad. God bless** u momo

    [ideation: parents/ attitude: negative TENACITY]

    [ideation: (it) parents/ attitude: negative PROPRIETY]

    [ideation: momo/ attitude: invoked positive JUDGEMENT]

    Convoke: Marshal (all u lazy parent) ↘ OUT GROUP COMMUNITY OF BAD PARENTS

    Finesse: Embellish (have to get involved, instead of) ↘ BAD PARENT BOND

    Moral evaluation: Evaluation

    De-Authorisation: Role Model

In these comments, evaluation is centred on the parents rather than on Momo and MORAL EVALUATION is also relied upon as a legitimation strategy for critiquing parents. CONVOKING affiliation is used to outgroup parents (BAD PARENT BOND) and generate moral panic about children using technology. Thus, negative judgement towards

parents and the de-legitimation of parents as role models take precedence over any issues regarding whether Momo is real or fake.

Moralisers were also outright critics of YouTube as a video platform. In these comments, negative judgement was directed towards YouTube. YouTube was portrayed as a something that either needed to be fixed or shut down. An example of this is seen in the following comment:

27. I remember when **everything was ok** on YouTube until the momo challenge virus started to be in videos **scaring many children and making them hurt themselves but end up dying because of it!** I just want YouTube to **go back to normal.**

    [ideation: YouTube/ attitude: positive NORMALITY]

    [ideation: momo challenge virus/ attitude: negative PROPRIETY]

    [ideation: YouTube/ attitude: negative NORMALITY]

    Finesse: Distil (started to be, because) ↘ BAD YOUTUBE BOND

    Temper: Foster (everything, many, just) ↘ DANGEROUS MOMO BOND

    Mythopoesis: Moral Tale

In this comment, YouTube switches from positive NORMALITY to negative NORMALITY, with TEMPERING affiliation creating a DANGEROUS MOMO BOND. There is a negative REACTION towards the Momo Challenge as it is described as being a 'VIRUS'. Thus, the Momo Challenge is imagined as something that is infecting videos on YouTube and legitimated via MYTHOPOESIS legitimation strategies that focus on the morality of the issue, as if this is the work of hackers and that it cannot be detected by others. Overall comments like this emphasise negative judgement towards YouTube as a video platform and a BAD YOUTUBE BOND.

Overall, Moralisers used a lot of TEMPERING affiliation to highlight the evils of the internet and persuade other to share these same reactions. Bonds were directed towards 'bad' actors, such as parents, hackers, and YouTube. MYTHOPOESIS: MORAL TALE was the most common legitimation strategy used, and was distinguished by the analyst from the MYTHOPOESIS: CAUTIONARY TALE strategies of the Myth Spreaders by the language instead emphasising the morality of the story being told by the commenter.

7.2.1.7. Summary of Results

Figure 7-1 illustrates an overview of the different personae encountered in the Momo Challenge dataset by providing a bond cluster diagram (previously published in Inwood and Zappavigna, 2021) of the logogenetic negotiation of particular bonds in the comment feeds. This diagram is based on the tendency of personae to share different patterns of bonds and attendant couplings, mapped out below in terms of their relation to 'real/fake Momo' and 'dangerous/safe Momo' bonds. It should be noted that the bond cluster diagram is a qualitative representation of the data analysed rather than a quantitative representation such as a computational text categorisation task.

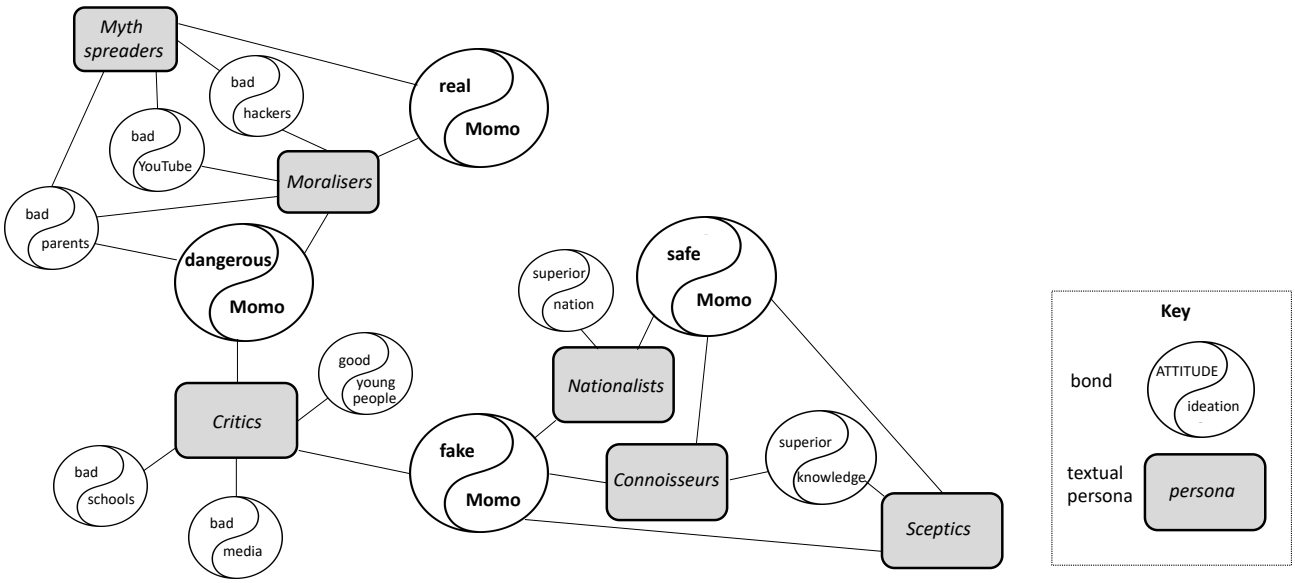


Figure 7-1 - Bond Cluster of Personae Discovered in the Momo Challenge Dataset

The bonds that cluster into each of the segments defined in this diagram give us a sense of the core worries (in the form of key bonds) that characterise the moral panic surrounding Momo, beyond the extent to which Momo is a hoax. Myth Spreaders and Moralisers aligned within the 'real Momo' and 'dangerous Momo' segment, negotiating bonds such as 'bad parents', 'bad YouTube' and 'bad hackers'. Critics involved themselves in the 'unsafe Momo' and 'fake Momo' segment, negotiating 'good young

people', and 'bad media' and 'bad schools' bonds. Nationalists, Connoisseurs, and Sceptics on the other hand operated within the 'safe Momo' and 'fake Momo' segment, negotiating 'superior nation' and 'superior knowledge' bonds. Some personae share a similar segment such as 'Myth Spreaders' and 'Moralisers', or 'Sceptics' and 'Connoisseurs'. In these cases, what distinguishes these personae from one another is the different linguistic features used to get to these bonds. This is illustrated in Table 7-1. For example, Myth Spreaders use more monoglossic and inconsistent language, while Moralisers are focused on negative judgemental language. Sceptics care more about asking questions while Connoisseurs prefer to state facts about internet culture.

This bond cluster diagram does not provide a hierarchy of which personae is the most or least truthful but rather maps out the different bond options available to each persona. In *Figure 7-2* the most common legitimation strategies for each persona and the most common bonds that correspond directly to these strategies are illustrated. Yin-Yang symbols that are in bold represent the bonds held by those sharing the 'momo is fake' sentiment: Connoisseurs, Nationalists, Sceptics, and Critics. The other yin-yang symbols represent the bonds held by personae sharing the 'momo is real' sentiment: Myth Spreaders and Moralisers. From this diagram we can see that AUTHORISATION is the most common legitimation strategy, particularly among personae that share the 'momo is fake' sentiment, with the dominant bonds being 'fake Momo', 'safe Momo', and 'superior nation' bonds. MYTHOPOESIS and MORAL EVALUATION were the more common legitimation strategies with personae that share the 'momo is real' sentiment, with these key bonds relating to 'dangerous Momo', 'real Momo', and 'bad YouTube'. Overall, this sort of diagram allows us to picture the relationship more clearly between key bonds and dominant legitimation strategies that relate to these bonds.

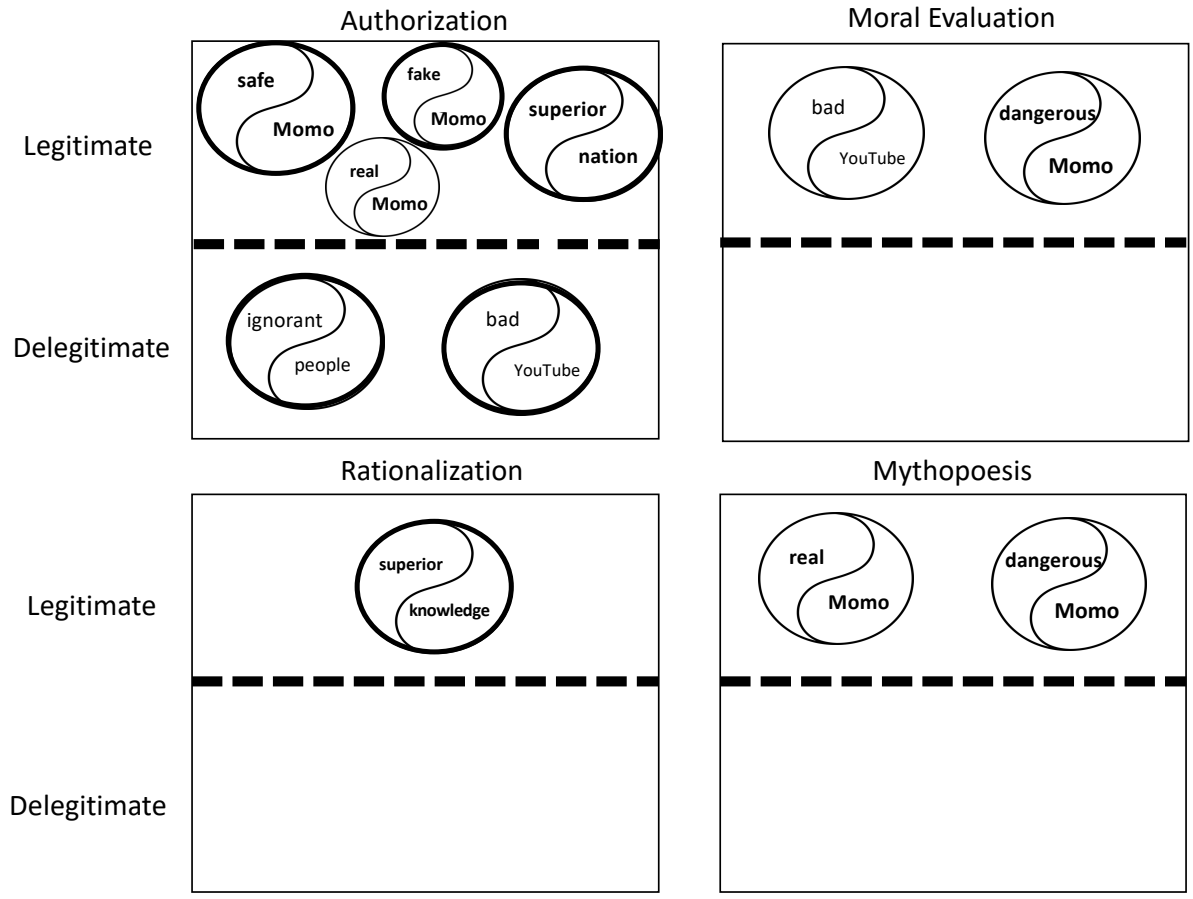


Figure 7-2 – Diagram of Key Bonds and Key (De)Legitimation Strategies

An overview of the frequency of the personae detected in the 750-comment corpus is provided in Table 7-1. This table provides an explanation of how the personae were identified based on groupings of shared linguistic features. Whilst this study represents qualitative research, Table 7-2 provides an overview of frequencies for key bonds and (de)legitimation strategies. This table adds justification to why certain bonds or legitimation strategies were focused upon in the examples given in the analysis section. As we can see 'fake momo' or 'real momo' bonds were most frequently represented in the comment corpus. Additionally, authorisation legitimation strategies (most notable TECHNOLOGICAL AUTHORITY) were frequently relied upon in both 'Momo is fake' and 'Momo is real' personae. Nonetheless, Myth Spreaders and Moralisers also relied upon MYTHOPOESIS as a core legitimation strategy in their comments. Thus, based on this case study, personae which used MYTHOPOESIS as a dominant legitimation strategy were more likely to be spreading false information about the Momo Challenge.

Personae	Frequency	Description of Personae	Key Linguistic Patterns	Most Common Legitimation Strategy
Connoisseurs	26.8%	Providing an 'insider' commentary on internet culture	FINESSING affiliation in relation to the proposition that Momo is an object not a person	AUTHORISATION: IMPERSONAL: TECHNOLOGICAL RATIONALISATION: THEORETICAL: EXPLANATION
Nationalists	4.1%	Emphasizing the importance of geographic location	CONVOKING affiliation to state that one's country has a superior understanding of Momo	AUTHORISATION: PERSONAL
Sceptics	8.9%	Questioning the veracity of the Momo challenge	Positive ATTITUDE targeted at evidence about Momo	AUTHORISATION: IMPERSONAL: TECHNOLOGICAL (DE)AUTHORISATION: IMPERSONAL: TECHNOLOGICAL RATIONALISATION: THEORETICAL: EXPLANATION
Critics	10.6%	Criticizing the media, established institutions, and older generations	Negative JUDGMENT targeted at those that spread hysteria and moral panic	AUTHORISATION: IMPERSONAL: TECHNOLOGICAL (DE)AUTHORISATION: IMPERSONAL: TECHNOLOGICAL MORAL EVALUATION: EVALUATION

Myth Spreaders	27%	Spreading stories about apparent dangerous behaviour linked to the Momo challenge	More monoglossic language, inconsistent statements, and TEMPERING affiliation to emphasize claims	AUTHORISATION: IMPERSONAL: TECHNOLOGICAL MYTHOPOESIS: CAUTIONARY TALE
Moralisers	22.4%	Policing the behaviour of parents and hackers, and platforms such as YouTube	Negative JUDGMENTAL evaluation towards the collective target and TEMPERING affiliation to emphasize claims	MORAL EVALUATION: EVALUATION MYTHOPOESIS: MORAL TALE

Table 7-1 - Overview of Personae Discovered

	Connoisseurs	Nationalists	Sceptics	Critics	Myth Spreaders	Moralisers
Number of Comments in Dataset	201	31	67	80	203	168
Most Common Bond 1	Fake momo bond (n=151)	Fake momo bond (n=28)	Fake momo bond (n=54)	Fake momo bond (n=74)	Real momo bond (n=159)	Bad hackers or YouTubers bond (n=95)
Most Common Bond 2	Superior knowledge bond (n=88)	Superior nation bond (n=25)	Superior knowledge bond (n=33)	Bad YouTube or media bond (n=34)	Dangerous momo bond (n=147)	Dangerous momo bond (n=86)
Most Common Bond 3	Safe momo bond (n=9)	Safe momo bond (n=3)	Ignorant people bond (n=2)	Bad schools bond (n=5)	Bad hackers or technology bond (n=42)	Real momo bond (n=77)
Most common Bond 4	Bad YouTube or media bond (n=2)	Superior knowledge bond (n=2)	Safe momo bond (n=1)	Good young people (n=2)	Bad parents bond (n=3)	Bad parent bond (n=19)
Authorisation	107	30	30	22	35	34
(De)Authorisation	22	2	12	27	2	9
Moral Evaluation	23	9	8	40	12	97
(De)Moral Evaluation	2	0	0	1	0	0
Rationalisation	96	2	19	8	3	4
(De)Rationalisation	0	0	0	14	0	0
Mythopoesis	20	5	3	16	187	82
(De)Mythopoesis	1	0	1	0	0	0

Table 7-2 – Overview of Frequencies for Key Bonds and (De)Legitimation Strategies in Dataset

7.2.2. Analysis of Replies: Dialogic Affiliation

Analysing the dialogic affiliation strategies used in the exchanges in the dataset offers a way of illustrating how the replies aligned or dis-aligned with the initiating comments. In this section we will consider whether there were any patterns in the dialogic affiliation strategies in the replies characterising particular personae uncovered in the initiating comments. This is useful for further understanding the interaction between users' when they bond around internet hoaxes, and whether repliers represent any of the key personae discovered in analysis of the initiating comments. Overall, replies were much shorter than the initiating comments and included less instances of appraisal. For the purposes of this analysis, replies to 'Momo is real' and replies to 'Momo is fake' comments will be analysed separately. However, overall, both 'Momo is real' and 'Momo is fake' comments received more supporting dialogic affiliation than rejecting dialogic affiliation (frequencies are provided in Table 7-3).

Initiating comments that received only one or two replies tended to express supporting dialogic affiliation, whilst initiating comments that received many replies expressed a mix of supporting and rejecting affiliation strategies. This accords with research by Thelwall et al. (2012) detailing how negativity can drive initiating comments that receive many replies, whilst positive sentiment is more common in videos that attract fewer comments. But it should be noted that this research does not aim to quantitatively show initiating comment to reply ratios, rather the aim is to understand the patterns in language and affiliation strategies across the datasets. These patterns in terms of main ideational and attitudinal targets, affiliation strategies and personae, will now be explored.

7.2.2.1. Replies to 'Momo is Real' Comments

Replies to 'Momo is real' comments often involved commenters asking for more information, insulting the commenter, or positively evaluating the initiating comment. Examples from all the dialogic affiliation categories were identified in the dataset. Comments exemplifying each dialogic affiliation category will now be examined, as well

as how these replies exemplify the rhetorical strategies used by the personae identified in the initiating comment dataset, by adopting strategies such as negative judgemental language or a lack of attitudinal language being used altogether.

There were two main groups of users that would rally around a comment; those who used minimal evaluative language and did not correspond with a persona, and those that would evaluate Momo and correspond with the Moraliser personae. Firstly, when users RALLIED around the social bonds in a comment and used minimal attitudinal language, they did not align with a persona in the initiating comment dataset, due to the lack of evaluative language and rhetorical resources used in these replies. In these cases, the main target of evaluation was the comment or the initiating commenter:

28. (user) **Thank you** for this, more people need to share this info
    [ideation: commenter/ attitude: positive HAPPINESS]
    MANAGE: SUPPORT: WARRANT: RALLY ↘ HAPPY COMMENTER BOND

29. **Well said** (user)! Thank you for mentioning this **important** topic and providing some **education** to the comment section J
    [ideation: commenter/ attitude: positive CAPACITY]
    MANAGE: SUPPORT: WARRANT: RALLY ↘ GOOD COMMENTER BOND
    [ideation: topic/ attitude: positive VALUATION]
    MANAGE: SUPPORT: WARRANT: RALLY ↘ IMPORTANT TOPIC BOND

30. I learned something new! **thanks** man
    [ideation: comment/ attitude: positive VALUATION]
    MANAGE: SUPPORT: WARRANT: RALLY ↘ GOOD COMMENT BOND

In these examples, there is positive attitude directed towards the comment or commenter. '*Education*' is positively evaluated and is attached to a MOMO IS FAKE bond, but these replies don't give enough detail regarding further social bonds that are at stake, for example, bonds around the media or YouTube as bad. However, with replies where the main target of evaluation was "Momo", there were more rhetorical clues regarding what the users were essentially bonding around:

31. Definitely I have young cousins who are getting to the stage of using iPads I **wouldn't want anything like <u>that</u> happening to them** fake or not

    [ideation: momo/ attitude: negative VALUATION]

    MANAGE: SUPPORT: WARRANT: RALLY ↘️ DANGEROUS MOMO BOND

32. I know <u>it's</u> **crazy**! Bunch of <u>trolls and pranksters</u> **taking it too far!**

    [ideation: it's (momo)/ attitude: negative REACTION]

    [ideation: trolls and pranksters/ attitude: negative PROPRIETY]

    MANAGE: SUPPORT: WARRANT: RALLY ↘️ DANGEROUS TROLLS BOND

33. read on the news here that police officers have seen the <u>videos</u> they are **warning** about.

    [ideation: videos/ attitude: negative PROPRIETY]

    MANAGE: SUPPORT: WARRANT: RALLY ↘️ DANGEROUS MOMO BOND

In these examples, users are either negatively judging Momo and hackers, or spreading a myth. In the first two examples (31 and 32) there is negative attitude directed towards Momo. Similar to the Moralisers generating a BAD HACKER BOND, in example 32, '*trolls and pranksters*' are also associated with negative PROPRIETY. In the example 33, Momo videos are associated with negative PROPRIETY and legitimised by the commenter referring to the news and police officers as trusted sources. The focus on negative judgement corresponds to the Moraliser persona.

The Myth spreader personae was evident in replies that ADJUSTED the coupling in the initiating comment, typically adding extra information to the initiating comment, with the user stating that they had discovered Momo:

34. **I GOT** A <u>VIDEO</u> OF MOMO AT 3AM IN A PEPPA VIDEO

    [ideation: video/ attitude: positive REACTION]

    MANAGE: SUPPORT: WARRANT: ADJUSTED ↘️ REAL VIDEO BOND

35. (user) It's **horrible** I also heard that <u>Momo</u> asked a kid to turn on the oven and don't tell anyone 😭 I think the kids house burned down 😭 😭

    [ideation: Momo/ attitude: negative PROPRIETY]

    MANAGE: SUPPORT: WARRANT: ADJUSTED ↘️ DANGEROUS MOMO BOND

36. Ik I've seen it last year on Reddit before. **Surprised** it's just now popping up.

[ideation: it (momo)/ attitude: negative INSECURITY]

R MANAGE: SUPPORT: WARRANT: ADJUSTED ↘ REAL MOMO BOND

In the first example (34), there is an invoked positive REACTION towards the YouTube video as the user is aligning with the sentiment that Momo is real. However, extra information is added by the user's exclamation that they have also discovered evidence of the Momo Challenge. In the second example (35), there is some negative attitude with Momo being linked to a kid's house burning down, resembling the language of a Myth Spreader. The third example (36) does not express a distinct moral judgement but rather surprise that Momo is suddenly reappearing.

Replies that DEFERRED from the REAL MOMO BOND in the initiating comment, also added extra information, but these replies did not directly align with the content of the initiating comment, rather adding sensationalist claims or insulting the commenter. These replies represented a mix of the Myth Spreader and Moraliser personae:

37. Just bumping this commenet, kid here in Philippines **commited suicide cause** of this momo challenge

[ideation: momo challenge/ attitude: negative PROPRIETY]

MANAGE: SUPPORT: DEFERRED ↘ DANGEROUS MOMO BOND

38. look it up in the internet. its already trending here in the Philippines. Some **serious sick** bastards at work.

[ideation: bastards (momo creators)/ attitude: negative PROPRIETY]

MANAGE: SUPPORT: DEFERRED ↘ EVIL PEOPLE BOND

39. (user) They don't sneak onto YouTube videos **dumbass**. The person making the video **intentionally puts** momo into the video

[ideation: commenter/ attitude: negative CAPACITY]

MANAGE: SUPPORT: DEFERRED ↘ IGNORANT PERSON BOND

[ideation: person making video/ attitude: negative PROPRIETY]

MANAGE: SUPPORT: DEFERRED ↘ EVIL PERSON BOND

In all these comments there is some aspect of negative PROPRIETY either towards Momo, or the creator of the Momo videos. In example 37, the user 'bumps' the comment by adding the false claim that a child committed suicide due to the Momo Challenge, characteristic of the Myth Spreader personae. In the next two examples (38 and 39), Momo creators shift to become central ideational targets and are morally negatively judged, aligning more with the Moraliser personae as they are policing the behaviour of hackers.

Replies that DISMISSED the social bonds in the initiating comment, explicitly denied the initiating comment's truthfulness claims and aligned with the Sceptic personae. These comments were presented as straightforward denial of the claims being made by the initiating commenter or phrased as a demand for evidence:

40. That is **fake** because it would be impossible for someone to **interfere with a tv broadcast**
    [ideation: comment/ attitude: negative VERACITY]
    MANAGE: REJECT: DISMISS ⬇ FALSE COMMENT BOND
    [ideation: people interfering/ attitude: negative CAPACITY]
    MANAGE: REJECT: DISMISS ⬇ INCAPABLE PERSON BOND

41. (user) **Gimme a link where this pops up in the middle of a video**.. link please
    [ideation: user/ attitude: negative VERACITY]
    MANAGE: REJECT: DISMISS ⬇ LYING USER BOND

42. No its not showing up in any videos its **fake/a hoax**
    [ideation: its (momo)/ attitude: negative VERACITY]
    MANAGE: REJECT: DISMISS ⬇ FAKE MOMO BOND

Therefore, these examples align with what a 'Sceptic' might write because they are encouraging evidence and alluding to the negative VERACITY of the YouTube video being commented on.

Replies that OPPOSED the social bonds in the initiating comment, aligned more with the rhetorical strategies used by the Sceptic or Connoisseur personae, and would present

an alternative coupling to what was being dismissed. These comments emphasised Momo or the commenter as the main ideational target:

43. I don't believe this happened. Troll. 'This maniac'? **What are <u>you</u> even saying**. And the 13 likes on this comment? What is that saying? <u>YouTube's</u> **problem** in a nutshell. Think, <u>people</u>. Be sceptical. [video creator], either expose or take down this comment and mine. **All part of the problem**. This is a **helpful** <u>video</u>, don't let it be polluted with anti-information.

    [ideation: you (momo commenter)/ attitude: negative VERACITY]

    MANAGE: REJECT: OPPOSE ⬇️ LYING PERSON BOND

    [ideation: YouTube/ attitude: negative PROPRIETY]

    MANAGE: REJECT: OPPOSE ⬇️ EVIL YOUTUBE BOND

    [ideation: people (video commenters)/ attitude: negative CAPACITY]

    MANAGE: REJECT: OPPOSE ⬇️ IGNORANT PEOPLE BOND

    [ideation: video/ attitude: positive VALUATION]

    MANAGE: REJECT: OPPOSE ⬇️ USEFUL VIDEO BOND

44. Google <u>momo</u> and it literally says it's a **hoax** and it's just a sculpture and the challenge are just hackers

    [ideation: momo/ attitude: negative VERACITY]

    MANAGE: REJECT: OPPOSE ⬇️ FAKE MOMO BOND

    [ideation: sculpture/ attitude: positive REACTION]

    MANAGE: REJECT: OPPOSE ⬇️ SAFE MOMO BOND

In the first comment (43), YouTube commenters broadly are opposed for their negative VERACITY, as well as YouTube as a video platform, for encouraging trolling behaviour. An alternative coupling is then proposed that the YouTube video has positive VALUATION, it should just avoid being polluted with 'anti-information'. This comment aligns with the Sceptic persona as the user directly doubts the information they are provided with from the initiating commenter. These Connoisseur personae was also evident in these sorts of comments shown by the second example (44) that aligns with the Connoisseur personae by encouraging others to '*google*' and understand internet culture better.

Lastly, replies that IGNORED the social bonds in the initiating comment entirely in the 'Momo is real' dataset, did not engage with the content of the initiating comment but would instead evaluate the initiating commenter's spelling:

45. If you have kids with that grammar that's **scarier** than momo. Your kid is probably a momo.
    [ideation: kids with bad grammar/ attitude: negative CAPACITY]
    IGNORE ⬇️ SAFE MOMO BOND

46. the fact you don't know the difference between "your" and "you're" could in fact mean you're the **fuckin idiot** duuuuuude
    [ideation: user/ attitude: negative CAPACITY]
    IGNORE ⬇️ IGNORANT PERSON BOND

47. You forgot the comma before your previous"and" comment. Better **work on your grammar** before becoming a nazi.
    [ideation: commenter/ attitude: negative CAPACITY]
    IGNORE ⬇️ IGNORANT PERSON BOND

These comments represent the commenter showing their superior knowledge and do not make any explicit commentary on the Momo Challenge. As there are no explicit bonds being shared about Momo, these comments do not align with the personae identified in the initiating comment dataset.

Overall, replies to 'Momo is real' comments predominately used negative evaluative language. However, there were a considerable number of comments that positively evaluated the commenter. A range of different dialogic strategies could be found within the dataset, revealing a mix of commenters aligning and dis-aligning with the initiating comment and resembling predominately Moraliser, Myth Spreader, or Sceptic personae.

7.2.2.2. Replies to 'Momo is Fake' Comments

Replies to 'Momo is fake' comments tended to use similar rhetorical strategies to the 'Momo is real' dataset in terms of negatively judging the commenter, or positively evaluating the initiating comment. Nonetheless, in this dataset there was more of an emphasis on people sharing facts and criticising the actions of parents, and an absence of people asking for more evidence. Examples from each of the different dialogic affiliation strategies available and how this can be considered in relation to personae, will now be shown.

When users RALLIED around the social bonds in an initiating comment, they often did not use evaluative language targeted towards specific individuals or institutions, instead simply agreeing with the initiating comment such as '*Exfrickingsactly*', '*exactly*', '*same*' and '*yeah I know*'. Thus, these short comments could not be aligned to the personae shown in the bond cluster diagram due to their lack of specific evaluations. When evaluative language was used, Momo and the verbiage of the initiating comment were again the main targets of evaluation:

48. **Thanks for the info** now I can tell my 5 year old brother he will be fine
    [ideation: comment/ attitude: positive VALUATION]
    MANAGE: SUPPORT: WARRANT: RALLY ↘ SAFE MOMO BOND

49. Lol of course momo is **fake**
    [ideation: momo/ attitude: negative VERACITY]
    MANAGE: SUPPORT: WARRANT: RALLY ↘ FAKE MOMO BOND

50. Exactly, I'm so lost as to **what the hell** everyones on about because as I'm doing research on this "momo" **crap** all I see is the **creepy** ladies image and people talking about it but not really saying what is is, just that it dares you to "do things" and people killing them selves and it's not an app it just "pops up" on things, so I'm confused on how there's a "challenge" on something that can't even be accessed? I'm calling **BS**
    [ideation: momo/ attitude: negative VERACITY]
    [ideation: momo image/ attitude: negative REACTION]

MANAGE: SUPPORT: WARRANT: RALLY ↘ FAKE MOMO BOND

Examples 48 and 49 both rally around the initiating comment but express this rallying via different evaluations. Whilst example 48 positively evaluates the comment, example 49 focuses more on reinforcing the fact that Momo is fake. Due to the shortness of these comments, they cannot be convincingly linked to a persona. Example 50 is much more detailed but still expresses the same evaluation that Momo is fake. However, by referring to '*research*' and questioning the assumptions behind the challenge, this example aligns with a Sceptic persona.

Replies that ADJUSTED the social bonds in initiating comment, extended the initiating comment by evaluating Momo as a meme or hoax, hence agreeing with the main sentiment that Momo is fake:

51. No one does. It **doesn't exist**. It's a joke on **naive** adults.
    [ideation: it (momo)/ attitude: negative VERACITY]
    MANAGE: SUPPORT: WARRANT: ADJUSTED ↘ FAKE MOMO BOND
    [ideation: adults/ attitude: negative CAPACITY]
    MANAGE: SUPPORT: WARRANT: ADJUSTED ↘ IGNORANT ADULTS BOND

52. Momo is a **social scare** that was originally a japanese art project that someone turned into a **worldwide global scare**
    [ideation: momo/ attitude: negative VALUATION]
    MANAGE: SUPPORT: WARRANT: ADJUSTED ↘ FAKE MOMO BOND
    [ideation: momo/ attitude: negative REACTION]
    MANAGE: SUPPORT: WARRANT: ADJUSTED ↘ SAFE MOMO BOND

53. I'm trying to convince my friend's its a **meme** but they're like "BUT ITS **SCARY** AFFZGNZJWJSJSNSJSE" Guess its not a meme in the UK, **lmao**
    [ideation: momo/ attitude: negative VERACITY]
    MANAGE: SUPPORT: WARRANT: ADJUSTED ↘ FAKE MOMO BOND
    [ideation: friends' reactions to momo/ attitude: negative DISINCLINATION]
    MANAGE: SUPPORT: WARRANT: ADJUSTED ↘ IGNORANT FRIEND BOND
    [ideation: momo in the UK/ attitude: negative REACTION]

MANAGE: SUPPORT: WARRANT: ADJUSTED ⬇️ FAKE MOMO BOND

These comments all express a negative attitude towards Momo and would align with a fake Momo bond. By referencing that Momo is a '*joke*', '*social scare*', or '*meme*' these comments would align with commenters representing the Connoisseur personae, as they have a more intricate understanding of internet culture.

Replies that DEFERRED the social bonds in the initiating comment, added extra information that was not related to the core evaluative claims of the initiating comment:

54. There are clips of Momo popping up on children's videos **instructing little kids to cut their wrists**, so it's become **relevant again**

    [ideation: Momo/ attitude: negative PROPRIETY]

    MANAGE: SUPPORT: DEFERRED ⬇️ DANGEROUS MOMO BOND

55. Just a rando fact: Momo was originally called "Mother Bird" but somehow ended up Momo like everyone knows her as. She was also destroyed so that "**she doesn't exist and the curse is gone.**"

    [ideation: Momo/ attitude: positive SECURITY]

    MANAGE: SUPPORT: WARRANT: ADJUSTED ⬇️ SAFE MOMO BOND

56. Yes momo **isn't real** but the trolls and pranksters **are**

    [ideation: momo/ attitude: negative VERACITY]

    MANAGE: SUPPORT: WARRANT: ADJUSTED ⬇️ FAKE MOMO BOND

    [ideation: trolls and pranksters/ attitude: positive CAPACITY]

    MANAGE: SUPPORT: WARRANT: ADJUSTED ⬇️ EVIL TROLLS BOND

These comments resemble a mix of different personae. Example 54 is a Myth Spreader as this comment is spreading the rumour that Momo is appearing in YouTube videos with minimal explicit attitudinal language. Example 55 refers to Momo's status as a statue that has been destroyed, aligning with a Connoisseur sharing information about internet culture. Example 56 aligns with a Moraliser who is concerned more about the dangers of 'trolls and pranksters' than questions about Momo's veracity.

There were a smaller number of replies that DISMISSED or OPPOSED the social bonds in the initiating comment, in contrast to the 'Momo is fake' dataset. With the DISMISS comments, Momo or hackers were the main ideational targets:

57. No <u>that</u> **stupid**

    [ideation: momo/ attitude: negative CAPACITY]

    MANAGE: REJECT: DISMISS ↘ REAL MOMO BOND

58. (user name) if <u>its</u> hoax how are 5 people dead related to the momo challenge and **parents going on the news** talking about it.

    [ideation: its (momo)/ attitude: positive VERACITY]

    MANAGE: REJECT: DISMISS ↘ REAL MOMO BOND

59. The Dublin police department tweeted that parents should **keep their children safe**, because <u>Momo</u> was **"Hacking" into kids cartoons**

    [ideation: momo/ attitude: negative PROPRIETY]

    MANAGE: REJECT: DISMISS ↘ EVIL MOMO BOND

Example 57 simply states that Momo is stupid and cannot be conclusively linked to a persona. Example 58 aligns again with a Myth spreader persona by trying to definitely claim that the Momo Challenge is linked to suicides without providing evidence or extensive evaluations. Example 59 DISMISSES the initiating comment by using a tweet by the 'Dublin police department' as evidence of the negative PROPRIETY of the Momo Challenge and aligns with a Moraliser persona.

Replies which OPPOSED the social bonds in the initiating comment would negatively evaluate Momo, parents or the YouTube platform, proposing an alternative coupling to the standard coupling of Momo being 'real'.

60. (user name) Hahaha, <u>that's</u> **funny**, I wish there was no reports, but there was an 11 year old that **committed suicide**, and a 13 found cutting her wrists, BC of the <u>momo challenge</u>

    [ideation: that's (previous comment)/ attitude: negative VERACITY]

    [ideation: report/ attitude: positive VERACITY]

    MANAGE: REJECT: OPPOSE ↘ REAL MOMO BOND

[ideation: momo/ attitude: negative PROPRIETY]

MANAGE: REJECT: OPPOSE ⬇️ REAL MOMO BOND

61. (user) I'm guessing some channels caught up on the band wagon and decided to freak kids out. It's entirely plausible that there are Momo videos on YouTube kids. YouTube kids is a **mess,** and it's the parents' **fault** for letting their child watch that **shit**.

    [ideation: some channels and YouTube kids/ attitude: negative PROPRIETY]

    MANAGE: REJECT: OPPOSE ⬇️ BAD CHANNELS BOND

    [ideation: parents/ attitude: negative CAPACITY]

    MANAGE: REJECT: OPPOSE ⬇️ BAD PARENTS BOND

62. It's **not just a statue**. Hackers used it to **get peoples; data/personal information**.

    [ideation: it's (momo)/ attitude: positive VERACITY]

    MANAGE: REJECT: OPPOSE ⬇️ REAL MOMO BOND

    [ideation: hackers/ attitude: negative PROPRIETY]

    MANAGE: REJECT: OPPOSE ⬇️ EVIL HACKERS BOND

Example 60 spreads the rumour that suicides are linked to the Momo Challenge, thus aligning with a Myth Spreader persona, and the alternative coupling of the report having positive VERACITY. Example 61 aligns with a Moraliser focusing more on negative evaluation of parents rather than the VERACITY of the Momo Challenge. Example 62 aligns again with a Moraliser, concerned more about the alternative coupling of a hacker having negative PROPRIETY than the VERACITY of the Momo Challenge.

Lastly, replies that IGNORED the social bonds in the initiating comment, rarely used evaluative language and were spam-like in their replies, not engaging at all with the initiating comment. In similarity, with 'Momo is fake' comments, these replies would also correct the spelling mistakes of the initiating comment writer:

63. Momo

64. wut

65. *there

66. I'm new here can I please get a like I know I'm **weird**

[ideation: commenter/ attitude: negative capacity]

IGNORE ↘ STRANGE COMMENTER BOND

67. Well,OoOf

These replies ignore the ideational targets in the initiating comments and do not correspond to a particular personae.

7.2.2.3. Summary of Comment Analysis

Overall, there were more similarities among the replies to the Momo is real and Momo is fake datasets, than differences. Commenters used a range of dialogic affiliation strategies. Although similar dialogic strategies were used, there was a difference in ideational targets chosen by 'Momo is real' vs. 'Momo is fake' supporters. 'Momo is real' comments focused on ideational targets regarding the emotional impact of Momo, whilst 'Momo is fake' comments focused on ideational targets regarding criticising parents and criticising the commenter for not doing their research. This is why the Sceptic personae was only featured in the 'Momo is real' dataset. To summarise the main personae discovered, 'Momo is real' replies aligned most with Connoisseurs, Sceptics, Myth Spreaders and Moralisers, whilst 'Momo is fake' aligned with Connoisseurs, Myth Spreaders and Moralisers. In all, replies were quite diverse in opinion, with a range of examples of commenters agreeing, disagreeing, and ignoring the initiating comment. A summary of the macro-level affiliation strategies in the replies dataset is provided in *Table 7-3*.

Type of Comment	Number of Initiating Comments with Replies	Number of Replies (Total)	Macro-Level Affiliation Strategies	
Momo is real	79	309	Agree	181
			Disagree	93
			Ignore	35
Momo is fake	79	198	Agree	122
			Disagree	45
			Ignore	31

Table 7-3 - Replies Analysis Dataset

7.3. Personae in the Notre Dame Fire Case Study

This section explores data from the Notre Dame Fire case study, in particular how the underlying social bonds proposed in white supremacist and conspiratorial discourse are discursively negotiated in the YouTube comment dataset. By conducting a close qualitative analysis of these comments about the Notre Dame Fire, a target of many xenophobic and conspiratorial claims, this section identifies a range of textual personae who respond to the conspiracy theories in the videos. It also explores the (de)legitimation used to create credible bonds. *The Methodology Chapter* outlines the sampling strategy used to collect 50 initiating comments from each of the 15 videos included in the dataset. The replies of each of these initiating comments were also analysed, with the findings presented in later in this chapter The main research questions for the comment analysis were:

i. What are the main ideational targets and attitudes that users' bond around when discussing conspiracies about the Notre Dame Fire on YouTube? How are these social bonds legitimated?

ii. By analysing the language of these comments, can a system of textual personae categories be developed? How does the notion of textual personae illuminate issues related to conspiratorial and white supremacist discourse?
iii. What patterns in language and affiliation strategies emerge when analysing replies to initiating comments?

7.3.1. Analysis of Initiating Comments: Communing Affiliation and Legitimation

The following results will show the different personae categories that were identified by the analysis focused on evaluative language and affiliation strategies. In addition, for each persona the most common legitimation strategies will be identified and how these strategies relate to persuasively bolstering each persona's key values.

7.3.1.1. Anti-Elitist

Anti-Elitists focused on depicting politicians and authorities as evil via negative JUDGEMENT and DEAUTHORISATION strategies. By elite, this means a select body of people who prefer to rule by 'elite system', namely those elected to office (politicians). DISTILLING was their most common affiliation strategy, used to definitively state that the fire was deliberate, as in 68 which claims that the fire was *'definitely a false narrative event'*, forming a *DELIBERATE FIRE BOND* realised by the coupling of the ideational target 'event' with negative VALUATION. DESIGNATING affiliation was also evident in order to name particular individuals. In Example 68 DESIGNATING affiliation is used to depict public officials as suspicious, with the ideational target of public officials being associated with negative VERACITY. This establishes a *LYING AUTHORITIES BOND*. This sentence is also crucial to the comment in terms of (de)legitimation strategies, as the assumption that officials are lying about the cause of the fire, delegitimises them as experts in the community. Thus, we can see here that the negative bonds are working in tandem with delegitimation strategies. Lastly, this comment ends with *"Piece of crap"* in quotation marks as an instance of EMBELLISHING affiliation because the statement is

being quoted, hence there is a voice attributed to it. Attitudinally, *"Piece of crap"* represents negative REACTION towards the event itself, thus forming a *FALSE EVENT BOND*, strengthening the initial claim that the Notre Dame Fire is a *'false narrative event'* and public officials are liars.

68. Definitely a **false** narrative event. **Not a fake event**, but public officials know who set this icon on fire, but for the sake religion of "peace" they **cover it up**. "**Piece of crap**" more like it.

    [ideation: public officials/attitude: negative VERACITY]

    [ideation: it (event)/attitude: negative REACTION}

    Finesse: Distil (definitely) ↘ DELIBERATE FIRE BOND

    Convoke: Designate (public officials, they) ↘ LYING AUTHORITIES BOND

    Finesse: Embellish ("piece of crap") ↘ FALSE EVENT BOND

    De-Authorisation: Commendation: Expert

Comments from Anti-Elitists all had in common negative JUDGEMENT towards elites such as politicians and authorities, and the specific delegitimation strategy of DE-AUTHORISATION: COMMENDATION: EXPERT. Bonds would overwhelmingly emphasise elites as *LYING* or *EVIL*, thus delegitimating their role in society as experts or role models. Compared to other personae categories (see *Table 7-5*), Anti-Elitists only had the one commonly occurring delegitimation strategy of DE-AUTHORISATION: COMMENDATION: EXPERT, and each comment contained a *LYING ELITES* or *EVIL ELITES* bonds, supplemented with a *SUSPICIOUS FIRE* or *DELIBERATE FIRE BOND*. Overall, this meant that Anti-Elitists comments were easiest to identify due to the smaller range of bonds and delegitimation strategies they would incorporate.

7.3.1.2. Anti-Media

The Anti-Media personae establishes the mainstream media (typically abbreviated as MSM) and social media platforms as evil through negative JUDGEMENT and DEAUTHORISATION strategies. Media typically refers to mass communication (Harper, n.d.), and thus Anti-Media persona target mainstream media in their evaluations. In similarity to the Anti-Elitist personae, DISTILLING affiliation was a common strategy for

making definitive statements, for example, in Example 69 although the commenter places emphasis on their subjective knowledge *'I knew'* and *'my mind'*, they express this definitively, refusing to acknowledge that they could be wrong. When they express that '*I knew it was arson'* arson has an invoked positive VALUATION because the commenter is positively evaluating their knowledge that the fire was deliberate, therefore establishing a *DELIBERATE FIRE BOND*. On the other hand, *'MSM'* is coded as invoked negative VERACITY because the mainstream media's reports will not change the commenter's mind, hence the *'MSM'* is an untrusted source of information. In considering the negative VERACITY of *'MSM'* and the DISTILLING affiliation of *'no'* as again this is said definitively, this delegitimates the MSM. DE-AUTHORISATION: CUSTOM: CONFORMITY has been chosen here because mainstream media is represented as an example of conformity, thus, to reject the mainstream media is to delegitimise conformity. In this example, we can see that delegitimation strategies supplement the bond that accuses the MSM of being an unreliable source of information.

69. **I knew** it was arson the moment I saw it on fire. No MSM report **will change my mind.**

[ideation: arson/ attitude: invoked positive VALUATION]

[ideation: MSM/ attitude: invoked negative VERACITY]

Finesse: Distil (I knew) ↘ DELIBERATE FIRE BOND

Finesse: Distil (no) ↘ FAKE MSM BOND

De-Authorisation: Custom: Conformity

In another comment (70) also presents a media company, in this case Buzzfeed, as an unreliable source and a target of invoked negative VERACITY. The company is depicted as hypocritical for criticising people making political use of the fire, while using the same strategy themselves. Semicolons are used in the comment as kind of projection marker, suggesting that what follows is Buzzfeed's action and opinion (with a similar function to projected speech in quotation). The parallelism in the '*also Buzzfeed:…'* is used to indicate the apparent hypocrisy and to enact a form of mockery. This also draws upon the common meme template of 'Me: X, Also Me: Y' where X represents a positive statement about oneself whilst Y contradicts the X statement. This usage, together with the CONVOKING affiliation seen in the collectivising *'our'* rallies the audience around a

HYPOCRITICAL MEDIA BOND. In terms of legitimation, the mockery acts in the service of DE-AUTHORISATION: COMMENDATION: EXPERT, with EXPERT rather than CONFORMITY as the sub-selection within this system because Buzzfeed is not explicitly depicted as mainstream media, and journalists can be regarded as experts in society due to the knowledge they bring to public discourse.

70. Buzzfeed: people are using this fire to attack their political enemies, also Buzzfeed: **let's use this fire to attack our political enemies.**

    [ideation: Buzzfeed/ attitude: invoked negative VERACITY]

    Convoke: Marshal (our) HYPOCRITICAL MEDIA BOND

    De-Authorisation: Commendation: Expert

The Anti-Media personae was quite similar to the Anti-Elitist personae in terms of affiliation strategies and the prominence of negative JUDGEMENT. Nonetheless, these categories of personae have been separated due to the difference in ideational targets and legitimation strategies – the *'MSM'* being the sole target of the Anti-Media personae, whilst the Anti-Elitist personae negative evaluated a range of people such as politicians, fire management authorities, and public officials. In addition, while Anti-Elitists would delegitimate experts predominately, Anti-Media personae would delegitimate both experts and societal conformity (i.e., those who adhere to mainstream news).

7.3.1.3. Inciter

To incite, means to provoke or to excite (Harper, n.d.). Inciters propagated suspicion regarding the cause of the fire, however, didn't specifically name a perpetrator as the cause. Instead, an object or anonymous entity such as the "*Deep State*" or "*Direct Energy Weapons*" were mentioned. In terms of legitimation, Inciters engaged primarily with one of three strategies: legitimatising their own narratives, legitimising technology as evidence, and delegitimising the words of experts. An example of Inciters focusing on a non-human entity is shown in Example **71.** In this example, *'burning'* is an invoked negative REACTION, strengthened by the TEMPERING affiliation of *'very'*. This forms a *SUSPICIOUS FIRE BOND* because it is emphasising the strange way the fire is burning. *'DEW*

Direct Energy Weapon' is treated with negative VALUATION. The DISTILLING affiliation of *'caused by'* asserts that the DEW caused the fire, hence this means a *BAD DEW BOND*. Although a person is mentioned in this comment as *'someone'*, they are unnamed, and the EMBELLISHING affiliation of *'could've'* portrays the commenter as uncertain about whether they actually saw the lasers. Both the *'someone'* as in the person who fired the lasers and the workers are portrayed with positive CAPACITY for their knowledge. With this in mind, it means a *SUSPICIOUS FIRE BOND* is formed. In the last sentence, lasers have an invoked negative VALUATION because they were fired at the Notre Dame. *'Fired at'* is an example of DISTILLING affiliation because it is stated definitively with no attributed voice. This forms a *BAD DEW* bond again because it associated DEW with causing the Notre Dame Fire. This last sentence is also crucial for legitimising Facebook as a form of evidence. The definitively stated phrase *'photo's on Facebook showed a laser'* (a DISTILLING AFFILIATION strategy) is legitimated as evidence for DEW being the cause of the Notre Dame Fire.

71. When I saw the first views from the fire, I immediately thought at the California fires, cause of the same kind of **profoundly way** of burning. These **earler fires were caused by** DEW Direct Energy Weapon. These lasers were fired by someone **whom was very aware** of the working hours of the workers. If they still were present, then someone **could've seen the lasers.** Some photo's on facebook showed a laser, **fired at the Notre Dame.**

    [ideation: burning/ attitude: invoked negative REACTION]

    [ideation: DEW Direct Energy Weapons/ attitude: negative VALUATION]

    [ideation: someone/ attitude: positive CAPACITY]

    [ideation: laser/ attitude: invoked negative VALUATION]

    Temper: MODULATE (same kind, very) ↘ SUSPICIOUS FIRE BOND

    Finesse: DISTIL (caused by, fired by, fired at, showed) ↘ BAD DEW BOND

    Finesse: EMBELLISH (could've) ↘ SUSPICIOUS PERSON BOND

    Legitimising: Authority: Technological

Example 72 is an example of the legitimation of a conspiracy theory. The phrase *'inside job'* means that the Notre Dame fire is associated with invoked negative VALUATION because it is suspicious, therefore forming a *SUSPICIOUS FIRE BOND*. In addition, TEMPERING affiliation is frequent in this comment with *'too many'* and *'too few'* emphasising the

suspicious circumstances of the fire. In terms of legitimation strategies, *'inside job'* evokes MYTHOPOESIS: SINGLE DETERMINATION because it attributes the Notre Dame Fire to a single possible narrative; that it was an inside job. Thus, this comment legitimises conspiracy theories by casting suspicion on the cause of the fire with MODULATING affiliation and invoked negative VALUATION.

72. **Inside job.** Too many coincidences, and too few fire hoses
    [ideation: (fire)/ attitude: invoked negative VALUATION]
    Temper: MODULATE (too many, too few) ↘ SUSPICIOUS FIRE BOND
    Legitimising: Mythopoesis: Single Determination

In Example **73,** experts and the Deep State are the core targets of negativity. *'Absurdly premature'* and *'absolute'*, as instances of MODULATING affiliation, draw attention to the ideational targets of *'denial of arson'* and *'Deep State'* via emphasising language. *'Denial of arson*' is an invoked negative VERACITY, forming a *LYING AUTHORITIES BOND* because it is criticising those who deny arson. *'Deep State'* is also an instance of negative PROPRIETY, as the organisation is evil for setting fire to Notre Dame, which means an *EVIL DEEP STATE BOND* is formed. As *'Deep State'* itself remains a vague entity (we don't know who belongs to this supposed organisation exactly), this ideational target fits the scope of the Inciter personae. In all, this comment is delegitimising experts by casting suspicion on those who deny arson quickly. Thus, the bonds analysis links again to the findings of the legitimation analysis.

73. The **absurdly premature** denial of arson is **absolute proof** that the Deep State **set fire to Notre Dame.**
    [ideation: denial of arson/ attitude: invoked negative VERACITY]
    [ideation: Deep State/ attitude: negative PROPRIETY]
    Temper: Modulate (absurdly premature) ↘ LYING AUTHORITIES BOND
    De-Authorisation: Commendation: Expert
    Temper: Modulate (absolute) ↘ EVIL DEEP STATE BOND

From these examples we can see that Inciters predominately direct negative JUDGEMENT towards non-specific entities, such as Direct Energy Weapons or the Deep State, and

predominately cast *SUSPICIOUS FIRE BONDS* or bonds stating that the Deep State or specific weapons were evil. DISTILLING affiliation and TEMPERING affiliation were the most affiliation strategies used in order to definitively state a claim or to draw attention to a claim. Compared to the other personae that shared false information about the fire, Inciters used a wider array of (de)legitimation strategies but were particularly focused on the legitimisation of conspiracy theories as credible sources of information.

7.3.1.4. White Supremacist

White Supremacists targeted a particular religion via negative JUDGEMENT and indicated a clash between different religions or cultures. White supremacists refer to those who believe in supreme authority, being supporters of the doctrine that white people are superior to other people and should have greater power and status (Harper, n.d.). This persona delegitimated the voices of governments and authorities, and instead legitimated YouTube videos and screenshots as credible sources of information. In Example 74 Muslim immigrants are depicted with negative PROPRIETY for causing the Notre Dame Fire and being *'illegal'*, forming an *EVIL MUSLIM BOND*. The CONVOKING affiliation of *'we'* means that the commenter's ambient audience includes those with a similar opinion. In addition, the phrase *'we will never know the truth'* delegitimises experts as it assumes that the experts or authorities are hiding something and are not telling the truth. In contrast, the phrase *'I bet my life's savings'* legitimises the commenter's prediction as superior to that of the experts. Thus, this comment uses various legitimation strategies to support the commenter's xenophobic remarks.

74. We will never know the truth but I bet my life's savings an **illegal** Muslim immigrant **started the fire.**

    [ideation: Muslim immigrant/ attitude: negative PROPRIETY]

    Convoke: Marshal (we) ↘ EVIL MUSLIM BOND

    De-Authorisation: Commendation: Expert

    Rationalisation: Theoretical: Prediction

In Example 75 references to technology are used by the commenter as evidence supporting their claim that Muslims are evil. Muslims and *'Muslim immigrants'* are depicted with negative PROPRIETY for celebrating the fire and for causing the fire. This establishes an *EVIL MUSLIMS BOND* throughout the comment. The TEMPERING affiliation of words such as *'sure'*, *'several'* and *'absolutely'* emphasises the claim that Muslims are evil. In addition, the DISTILLING affiliation of *'sure were'* and *'did it'* definitively states that Muslims celebrated and caused the fire. The legitimation analysis adds another layer of meaning to this analysis by emphasising how technology is used to create authority. The references to *'several videos circulating online'* and *'Facebook'* is used to legitimate the claims that Muslims celebrated and caused the fire, again leading to an *EVIL MUSLIMS BOND*.

75. Muslims **sure were in celebration** during the blaze. There's several videos circulating online that show them **screaming allahu akbar** at the fire. **They laugh about it on Facebook.** They **did it absolutely.**

    [ideation: Muslims & they/ attitude: negative PROPRIETY]

    [ideation: Muslim immigrant/ attitude: negative PROPRIETY]

    Temper: Modulate (sure, several, absolutely) ↘ EVIL MUSLIMS BOND

    Finesse: Distil (sure were, did it) ↘ EVIL MUSLIMS BOND

    Authorisation: Impersonal: Technological

These examples illustrate how White Supremacists are preoccupied with negatively judging non-western religions and cultures and setting up a dichotomy between the west versus another religion or culture. While there were some instances of negative judgment towards Jews, Muslims were predominately attacked by White Supremacists (see *Table 7-5*). In some cases, bonds moved away from concern about the fire to focus on anti-immigrant and white supremacist discourses by directing hatred to non-western cultures. In terms of legitimation strategies, White Supremacists in similarity to other personae sharing false information relied on videos and screenshots as credible evidence, whilst delegitimising the voices of politicians and experts.

7.3.1.5. Confrontationist

Confrontationists call out a video for being false or call out a YouTuber for lying. Comments by Confrontationists tended to be quite short and would not explain the reasoning behind why the video was false. To confront means to bring "two parties face to face" for examination and discovery of the truth (Harper, n.d.). Politically, confrontation is also associated with a state of political tension and the coming of countries face to face, with the word gaining particular prominence during the 1963 America versus Russia state of cold war (Harper, n.d.). The word 'confront' has negative connotations as aggressive and is marked by an adversarial approach. Confrontationists create this dichotomy when directly calling out a YouTuber for spreading false information, focusing upon negative evaluations.

In contrast to the personae sharing false information, Confrontationists focused on delegitimating strategies, usually delegitimating the sources of evidence other YouTubers used to support their claims. For example, Confrontationists delegitimated the videos other personae tried to use as evidence for their claims, delegitimated the voices of YouTubers, and delegitimated the conspiracy theories set out by YouTubers. In Example 76 we can see a very short comment by a Confrontationist that directly calls out the YouTube video for being false. The word *'clickbait'* and the emoticon *':o('* resembling a sad face contribute to the negative VALUATION of the title and video. This creates a *FAKE VIDEO BOND* because it is criticising the title of the video for being misleading and also invoking that the entire video itself is misleading. *'Another'* and the emoticon *':o('* are instances of MODULATING affiliation because these features strengthen the negative VALUATION of the comment. In terms of legitimation strategies, this comment is de-authorising the validity of the video by stating that it is clickbait.

76. Another **clickbait** title. **:o(**

[ideation: title (video)/ attitude: negative VALUATION]

Temper: Modulate (another, :o( ) ↘ FAKE VIDEO BOND

De-Authorisation: Impersonal: Technological

Example 77 is another short comment that directly calls out YouTubers. *'You Guys'* is a CONVOKING strategy for directly addressing the YouTubers. It is associated with negative VERACITY as the YouTubers are accused of being misleading, therefore establishing a *LYING YOUTUBER BOND*. *'Narrative'* is evaluated as negative VALUATION for its falsity, forming a *BAD NARRATIVE BOND*. The *'absolutely'* as a form of MODULATING affiliation intensifies the claim that the narrative is false. In terms of (de)legitimation strategies, there are two instances of delegitimating occurring in this comment. The first is PERSONAL DE-AUTHORISATION, as the commenter degrades the status of the YouTuber by stating that they are *'full of shit'*. In the second occurrence, the narrative itself is delegitimised because it *'makes absolutely no sense'*. In this case, the (de)legitimation strategies align with the two key bonds of *LYING YOUTUBER* and *BAD NARRATIVE*.

77. You guys are **full of shit**, the narrative you're pushing **makes absolutely no sense.**

[ideation: guys/ attitude: negative VERACITY]

[ideation: narrative/ attitude: negative VALUATION]

CONVOKE: DESIGNATE (you guys) ↘ LYING YOUTUBER BOND

TEMPER: MODULATE (absolutely) ↘ BAD NARRATIVE BOND

De-Authorisation: Personal

De-Mythopoesis: Single Determination

Overall, the discourse of Confrontationists is characterised by 'calling out' the YouTuber who created the video or the people who spread conspiracy theories. In order to evoke this 'calling out' CONVOKING affiliation was the most widely used resource. Bonds were centred around the ignorance or lying behaviour of the YouTuber, rather than discussion about the Notre Dame fire as a specific event. The delegitimation strategies evident in Confrontationist comments also supported the claim that Confrontationists are focused on 'calling out' a video or YouTuber as false rather than explaining why the video or YouTuber is false.

7.3.1.6. Educator

In contrast to the strategy of the Confrontationist, Educators explained why the content of a YouTube video is false, rather than simply criticising a YouTube video for being false. These comments featured multiple bonds. Whilst the videos and screenshots that YouTubers referred to was delegitimised by Educators, the legitimation of rational thinking and expert opinion was also emphasised. In Example 78 the *'Allah Akbar'* recording is treated with negative VALUATION because it is fake as authorities have already determined that the fire was an accident, therefore forming a *MISLEADING VIDEO BOND*. The phrase *'the French police authorities already said'* is legitimating the words of experts and is written as a form of DISTILLING affiliation because it does not entertain other possibilities. In the second part of the clause, the only evaluation occurs regarding *'everybody'* which is an example of negative CAPACITY because people are ignorant enough to *'jump to conspiracies'*, thus this means that an *IGNORANT PEOPLE BOND* is formed. *'Not everything is a freaking conspiracy'* de-legitimises the voice of the YouTuber and is an example of MODULATING affiliation because the *'freaking'* intensifies the claim. Another instance of legitimation is in the phrase *'accidents can happen'* as this provides an explanation for the fire via the highlighting of rational claims based on expert opinion as opposed to conspiratorial claims. It total, this comment uses multiple legitimation strategies and bonds to explain why the YouTuber is incorrect.

78. **Fake news** Allah Akbar come on now the French police authorities already said that it was electrical short that caused the fire in the area where they were doing renovation accidents can happen why does everybody got to **jump to conspiracies not everything is a freaking conspiracy.**

    [ideation: Allah Akbar recording/ attitude: negative VALUATION]

    [ideation: everybody/ attitude: negative CAPACITY]

    FINESSE: DISTIL (said) ↘ MISLEADING VIDEO BOND

    TEMPER: MODULATE (freaking) ↘ IGNORANT PEOPLE BOND

    Authorisation: Commendation: Expert

    Rationalisation: Theoretical: Explanation

    De-Mythopoesis: Single determination

In Example 79 the *'allahu ackbar audio'* is discussed again but this time receives more prominence throughout the entire comment. The *'allahu ackbar audio'* or *'meme'* has invoked negative VALUATION because it is misleading, hence forming a *MISLEADING VIDEO BOND*. The DISTILLING affiliation of *'pretty sure'* and *'it's'* are definitive statements, contributing to the YouTuber's certainty that this is a misleading video. The delegitimating of the 'allahu ackbar audio' and 'meme' as evidence is the only (de)legitimation strategy to occur in this comment. Thus, this comment is formed around the delegitimation of the Allahu ackbar recording and an explanation for why it is misleading.

79. Pretty sure that "allahu ackbar" audio was **dubbed over fire footage.** It's practically a **meme**, dubbing "allabu ackbar" over any disaster and "fail".

    [ideation: Allahu ackbar audio/ attitude: invoked negative VALUATION]

    [ideation: it (Allahu ackbar audio)/ attitude: invoked negative VALUATION]

    FINESSE: DISTIL (pretty sure, it's) ↘️ MISLEADING VIDEO BOND

    De-Authorisation: Impersonal: Technological

Comments by Educators also differed from other personae due to the more frequent use of positive evaluation. For example, in Example 80 the ideational target is shifted away from xenophobic views (negative judgement) to focusing on the idea of most people being good citizens (positive judgement). As we can see *'people of Muslim background'* and '*Muslim people and the commenter'* are evaluated with positive PROPRIETY because they genuinely care about the fire occurring, hence forming *GOOD MUSLIM* and *GOOD PEOPLE* bonds. *'I can attest'* establishes that the commenter is making a definitive statement. *'Many were'* references a community, therefore establishing DESIGNATING affiliation. This comment ends on a negative evaluation, as the Notre Dame Fire is associated with negative REACTION for being *'a loss for entire mankind'* and intensified with MODULATING affiliation to create a *SAD FIRE BOND*. However overall, in this comment, more positive bonds are emphasised than negative bonds. In terms of legitimation strategies, the commenter's own personal voice is being legitimised – *'I am an Albanian of Muslim background'* and *'I can attest that at the place where I work'* also affords status to the commenter's own experience and background.

80. I am an Albanian of Muslim background. I can attest that at the place where I work, people (most of whom of Muslim background) gathered and **expressed concerns** about the fire. Many were with **tears in eyes**; me too. This is a **loss** for entire mankind.

[ideation: people of Muslim background/ attitude: positive PROPRIETY]

[ideation: Muslim people and user/ attitude: positive PROPRIETY]

[ideation: this (Notre Dame)/ attitude: negative REACTION]

FINESSE: DISTIL (I can attest) ↘ GOOD MUSLIM BOND

CONVOKE: DESIGNATE (many were) ↘ GOOD PEOPLE BOND

TEMPER: MODULATE (entire) ↘ SAD FIRE BOND

Legitimising: Authority: Personal

In general, comments by Educators differed from other personae due to a mix of different ideational targets, more instances of positive evaluation, and comments that featured more bonds and affiliation strategies in order to explain why a person was incorrect. Educator (de)legitimation strategies were also more diverse compared to other personae, with comments featuring multiple legitimation strategies occurring in the same comment, also being very common.

7.3.1.7. Sceptic

Sceptics made comments casting doubt on the YouTube video or asking for more clarity. This persona delegitimised the idea of using YouTube videos as evidence and legitimised rational thinking by questioning the veracity of the YouTube video. In Example 81 it can be assumed that the commenter is discussing the video when they state that *'I'm not sure what I'm seeing'*, therefore *'video'* has invoked negative VALUATION for being misleading, hence forming an *UNCLEAR VIDEO BOND*. The *'not sure what I'm seeing'* complements this bond by delegitimating the video as a form of evidence. Thus, this shows the common structure for a comment by Sceptics; a short comment asking for more clarity.

81. I'm **not sure what I'm seeing** to make a clear judgment.

[ideation: video/ attitude: invoked negative VALUATION]

Finesse: Embellish (note sure) ↘ UNCLEAR VIDEO BOND

De-Authorisation: Impersonal: Technological

In another instance in Example 82 we see a longer comment by a Sceptic. This comment starts with a question, another common rhetorical structure for Sceptics. *'Religious nut jobs'* is evaluated as invoked negative PROPRIETY for stating the possibility that religious fanatics could set a church on fire, therefore forming an *EVIL RELIGIOUS PEOPLE BOND*. *'Wouldn't put it past'* is an instance of EMBELLISHING affiliation because it is guessing that religious fanatics could cause a fire. In the next instance, conspiracy theory is evaluated as negative VALUATION because it is described as *'crazy'*, therefore forming a *BAD CONSPIRACY BOND*. The word *'just'* emphasises this evaluation, hence it is an example of MODULATING affiliation. In the last instance, *'contingencies'* is given negative VALUATION because it is unnecessary to apply contingencies, forming an *UNSUSPICIOUS FIRE BOND*. The phrase *'never multiply'* is an example of DISTILLING affiliation because it definitively states that unnecessary contingencies shouldn't be multiplied. In terms of (de)legitimation, there are two key strategies. The focus on *'evidence'* is an example of legitimising rationalisations, in the sense that proper explanations with verifiable evidence need to be provided in order for the commenter to believe in the YouTuber's claims. Secondly, there is a focus on delegitimising conspiracy theories: by negatively evaluating conspiracy theories this means that they are also delegitimised as credible stories. Although this example is a longer than average comment from a Sceptic, it manifests the same rhetorical structure of requesting evidence.

82. Evidence? Or just opinion with no evidence? Don't get me wrong, I wouldn't put it past other **religious nut jobs** to **pull a stunt like this** but I'll reserve opinion until I have some evidence to support my assertions. Otherwise it is just a **crazy** conspiracy theory. Btw, look up Occam's razor and it basic principle. Never multiply unnecessary contingencies.

    [ideation: religious nut jobs/ attitude: invoked negative PROPRIETY]

    [ideation: conspiracy theory/ attitude: negative VALUATION]

    [ideation: contingencies/ attitude: negative VALUATION]

    FINESSE: EMBELLISH (wouldn't put it past) ↘ EVIL RELIGIOUS PEOPLE BOND

    TEMPER: MODULATE (just) ↘ BAD CONSPIRACY BOND

    FINESSE: DISTIL (never multiply) ↘ UNSUSPICIOUS FIRE BOND

Rationalisation: Theoretical: Explanation

De-Mythopoesis: Single Determination

In terms of linguistic strategies, the Sceptic personae is distinguished by shorter comments, questions, and EMBELLISHING affiliation that casts doubt or opens up the range of different possibilities. Sceptics most commonly delegitimised the videos of YouTubers as evidence, and legitimised rational language in terms of discussing what counts as evidence and what is scientifically possible.

7.3.1.8. Summary of Results

The personae identified in the above analysis can be mapped out in a bond cluster diagram in order to better understand how the key bonds in their discourse relate to each other (*Figure 7-3*). The bond cluster in *Figure 7-3* is illustrative of the key discoveries from a qualitative analysis; hence it should not be read as a quantitative social network analysis. The lines in the diagram represent the multiple bonds that the personae in the dataset held, as well as the bonds that were shared across different personae. The most prominent bonds shared among several personae are shown in bold, whilst bonds that are specifically related to one persona are depicted at a smaller scale.

In the bond cluster diagram, we can see that there are two core networks (interconnected structures rather than the SFL system network) in terms of the connections among bonds: the first network consists of the 'Anti-Elitist', 'Anti-Media', 'White Supremacist' and 'Inciter' personae (shown in dark grey). This network represents personae that propagated conspiratorial information about the Notre Dame Fire. In the second network we can see the 'Sceptics', 'Confrontationists' and 'Educators' personae (shown in light grey). This network represents personae that propagated anti-conspiratorial information about the Notre Dame Fire. Thus, this diagram is useful for showing the 'macro-persona' categories that exist in terms of grouping personae, as well as how each persona is unique in the particular bonds that they most commonly represent. The Notre Dame case study featured two distinct networks of macro-personae with no crossovers in core bonds. This shows the polarised nature of YouTube

comment threads and suggests how the absence of shared bonds can result in animosity as there is no common social bond (either positively or negatively evaluated) to serve as a basis of discussion and affiliation.

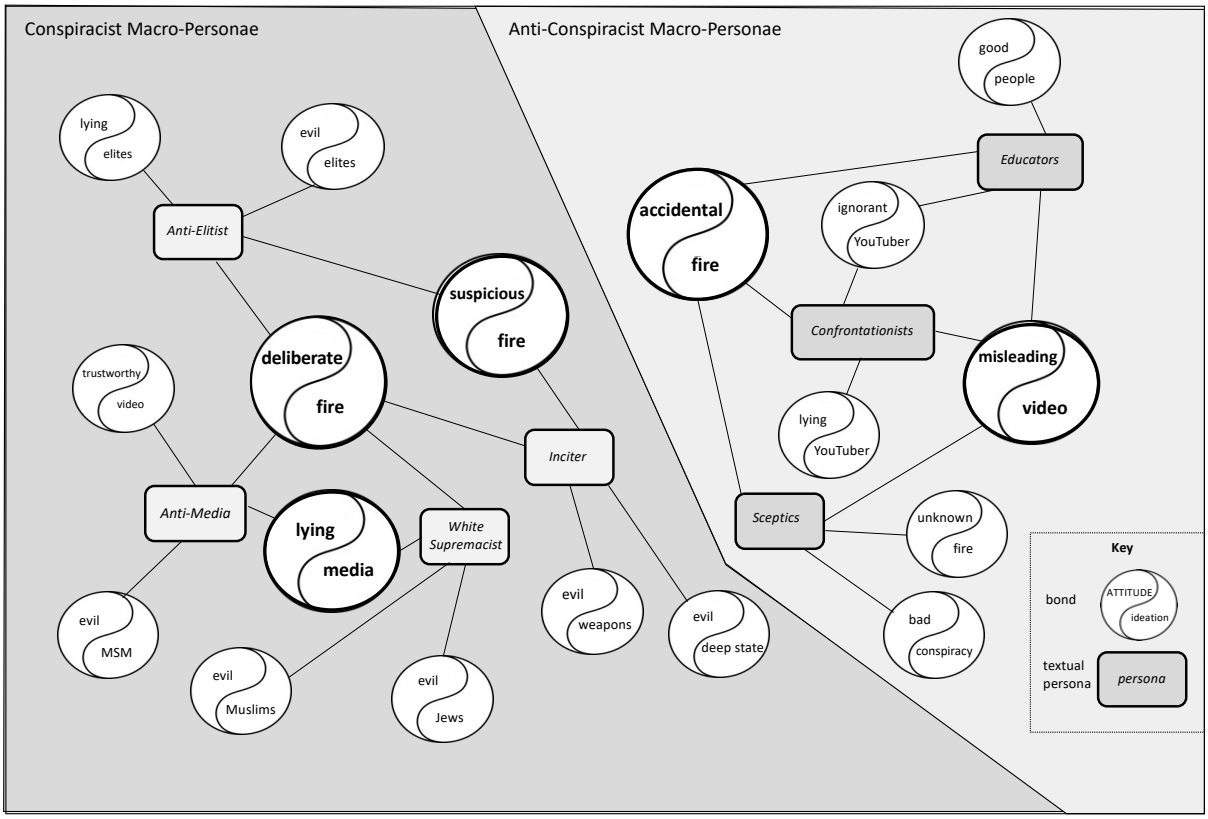


Figure 7-3 - Bond Cluster Diagram of Key Personae

In addition, another diagram has been created in order to illustrate the relationship between the key bonds and key (de)legitimation strategies discussed in the results. The diagram shown in Figure 7-4 illustrates the most common legitimation strategies for each persona and the most common bonds that corresponded directly to these strategies. The diagram only considers the four main categories of legitimation and is divided into legitimate and delegitimate sub-categories. Yin-Yang symbols in bold represent the bonds that were held by those sharing truthful information about the fire i.e. Sceptics, Confrontationists and Educators. The other Yin-Yang symbols represent the bonds held by personae sharing false information about the fire i.e. Anti-Elitists, Anti-Media, White Supremacists, and Inciters.

Based on the placement of the Ying-Yang symbols in the diagram we can see that de-authorisation was the most common (de)legitimation strategy according to personae sharing truthful or false information about the fire. However, the sub-categories differ – whilst Anti-Media and Anti-Elitists were deauthorising authorities and the MSM; Educators, Sceptics, and Confrontationists were deauthorising technology and individuals. Within the authorising category, Educators were authorising experts, whilst White Supremacists and Inciters were authorising technology. Thus, although this diagram provides an overview of the most common bonds and (de)legitimation groupings, the more delicate categories of (de)legitimation need to be explored in order to fully understand the discrepancies in legitimation strategies of personae sharing anti-conspiratorial information about the via versus personae sharing conspiratorial information about the fire.

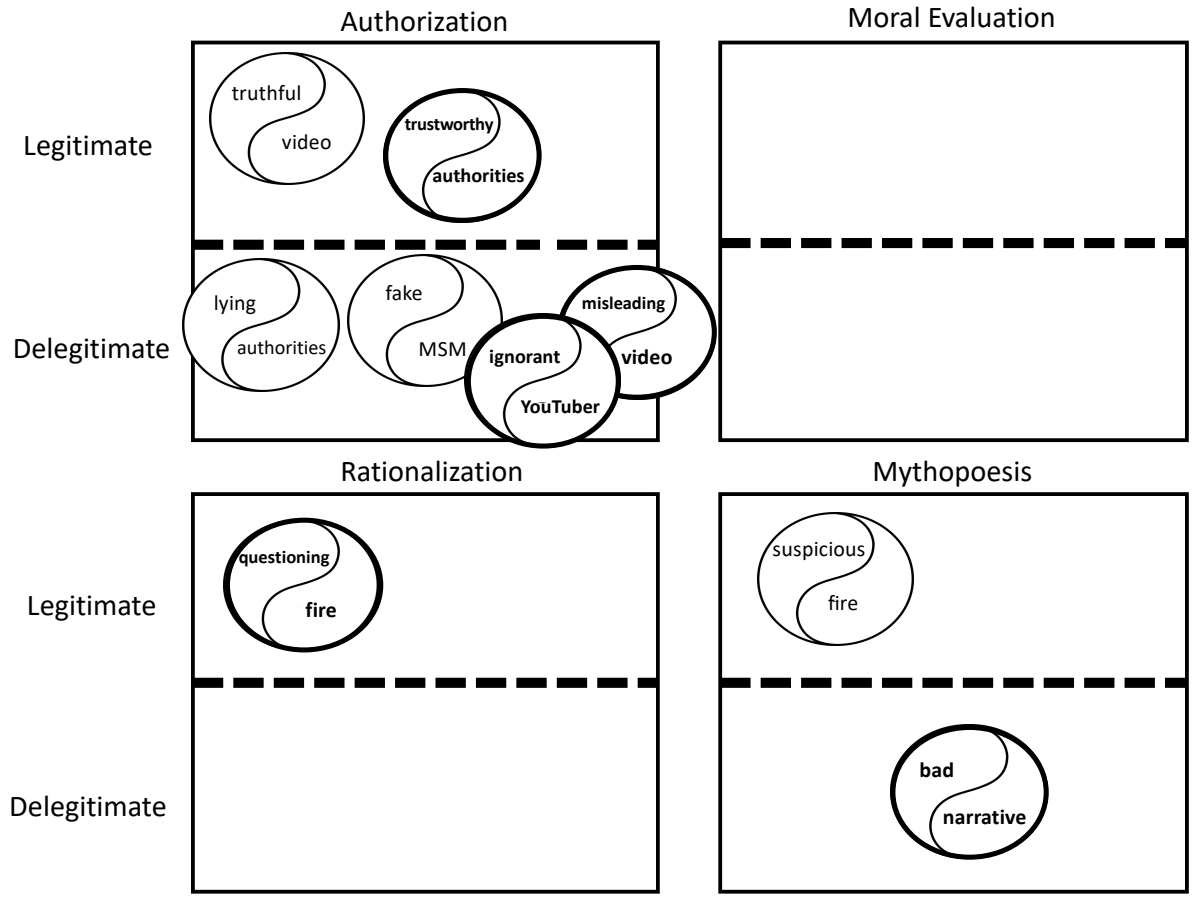


Figure 7-4 – Diagram of Key Bonds and Key (De)Legitimation Strategies

A summary of the key personae identified, and the key linguistic patterns and legitimation strategies for each persona is shown in *Table 7-4*. Whilst this study is qualitative, in *Table 7-5* an overview is provided of frequencies for key bonds and (de)legitimation strategies. The purpose of this table is to highlight the four most common bonds for each persona and how frequently they appeared in the dataset, as well as how frequently the four overarching categories of (de)legitimation appeared in the dataset. This table adds justification to the study's focus on explaining certain bonds or (de)legitimation strategies over others (with those being focused on highlighted in grey). It should also be noted that multiple bonds and multiple (de)legitimation strategies could appear in the same comment. Although some personae, such as the Inciter, most commonly featured one or two of the same bonds, other personae, such as the Educator or Confrontationist, tabled a more diverse array of bonds. This is why Educators and Confrontationists have lower frequencies in the dataset. Overall, these tables illustrate the scope of data that was analysed and why it is important.

Personae	Frequency in Dataset	Description of Personae	Key Linguistic Patterns	Key Legitimation Strategies
Anti-Elitist	11%	Focusing on the evil of authorities and politicians	Negative ideational target – authorities, DISTILLING AND DESIGNATING AFFILIATION	(DE)AUTHORIZATION: COMMENDATION: EXPERT
Anti-Media	4%	Focusing on mainstream media and social media platforms as evil	Negative ideational target – mainstream media, YouTube, DISTILLING AFFILIATION	(DE)AUTHORIZATION: COMMENDATION: EXPERT (DE)AUTHORIZATION: CUSTOM: CONFORMITY
White Supremacist	12%	Focusing on targeting a specific religion and stating one religion is superior to another	Negative ideational target – religion, contrasting two cultures, DISTILLING AND MODULATING AFFILIATION	(DE)AUTHORIZATION: COMMENDATION: EXPERT AUTHORIZATION: IMPERSONAL: TECHNOLOGICAL
Inciter	23%	Creating suspicion regarding the cause	Suspicious and deliberate suspicious	MYTHOPOESIS: SINGLE DETERMINATION

		of the fire but not specifically naming a person as the cause	fire bonds with no extra targets of evaluation, unnamed ideational target or non-human ideational target, TEMPERING AFFILIATION	AUTHORIZATION: IMPERSONAL: TECHNOLOGICAL (DE)AUTHORIZATION: COMMENDATION: EXPERT
Sceptic	4%	Expressing uncertainty about the video or information they are provided with	Only one target of evaluation, phrased as a question or embellishing the coupling	(DE)AUTHORIZATION: IMPERSONAL: TECHNOLOGICAL RATIONALIZATION: THEORETICAL
Educator	23%	Educating readers about why the video is false	Multiple targets of evaluation, More positive judgment – people as good	(DE)AUTHORIZATION: IMPERSONAL: TECHNOLOGICAL AUTHORIZATION: COMMENDATION: EXPERT RATIONALIZATION: THEORETICAL
Confrontationist	23%	Calling out something or someone as stupid/false, but not explaining why	Only one target of evaluation, negative ideational target is the YouTuber or video, CONVOKING AFFILIATION	(DE)AUTHORIZATION: IMPERSONAL: TECHNOLOGICAL (DE)AUTHORIZATION: PERSONAL: INDIVIDUAL (DE)MYTHOPOESIS: SINGLE DETERMINATION

Table 7-4 - Summary of Results

	Anti-Elitist	Anti-Media	White Supremacist	Inciter	Sceptic	Educator	Confrontationist
Number of Comments in Dataset	85	31	88	171	29	173	173
Most Common Bond 1	Evil elites bond (n=50)	Lying Media bond (n=27)	Evil Muslims bond (n=80)	Suspicious fire bond (n=128)	Misleading video bond (n=12)	Misleading video bond (n=60)	Ignorant YouTuber bond (n=56)
Most Common Bond 2	Lying elites bond (n=44)	Evil MSM bond (n=11)	Deliberate fire bond (n=22)	Deliberate fire bond (n=33)	Accidental fire bond (n=9)	Ignorant YouTuber bond (n=58)	Lying YouTuber bond (n=46)
Most Common Bond 3	Suspicious fire bond (n=27)	Deliberate fire bond (n=10)	Evil Jews bond (n=9)	Evil Deep State bond (n=12)	Unknown fire bond (n=4)	Accidental fire bond (n=52)	Accidental fire bond (n=12)
Most Common Bond 4	Deliberate fire bond (n=14)	Trustworthy video bond (n=6)	Lying media bond (n=9)	Evil weapons bond (n=11)	Bad conspiracy bond (n=3)	Good people bond (n=36)	Misleading video bond (n=15)
Authorization	14	8	24	57	7	57	15
(De)Authorization	74	29	42	49	12	91	102
Moral Evaluation	0	0	3	2	0	2	4

(De)Moral Evaluation	2	0	15	2	0	5	5
Rationalization	5	0	10	16	11	49	17
(De)Rationalization	0	0	0	2	0	15	7
Mythopoesis	8	5	15	72	0	1	0
(De)Mythopoesis	1	0	2	1	4	24	45

Table 7-5 - Overview of Frequencies for Key Bonds and (De)Legitimation Strategies in Dataset

7.3.2. Analysis of Replies: Dialogic Affiliation

This section will outline the key findings arising from analysing the replies to the initiating comments from the Notre Dame fire videos. This analysis focused on dialogic affiliation – a framework for analysing how replies aligned or dis-aligned with the initiating comments and the main social bonds used in order to achieve this aim (see *The Methodology Chapter* for an explanation of the dialogic affiliation system). Additionally, this analysis will consider how the replies align with the personae that were discovered in the primary comment dataset, in terms of whether the repliers adopt the same personae as the initiating commenters, or if they adopt different linguistic strategies to express their identity.

Replies to conspiratorial comments typically reinforced the view of the initiating commenter that the fire was deliberate or suspicious, with slight adjustments to the main ideational target – the specific group of people blamed for the fire. In comparison, replies to anti-conspiratorial comments, typically disagreed with the initiating commenter, changing the attitudinal and ideational target completely. In addition, conspiratorial comments received more engagement in comparison to anti-conspiratorial comments. These findings are not particularly surprising when

considering the video dataset which is conspiratorial videos about the Notre Dame Fire. The people engaging with these videos, already more likely share the same opinions as the YouTuber, as they belong to the same conspiratorial, white supremacist or far-right communities. However, what is interesting about this analysis is the way social bonds are constructed and the different textual personae that emerge. These findings will now be explored, by explaining each different dialogic affiliation strategy used in the dataset.

7.3.2.1. Replies to Conspiratorial Comments

Replies to conspiratorial comments typically aligned with the attitudes in the initiating comment, featuring xenophobic and anti-elitist values. ADJUSTING the social bonds in the comment was the main dialogic affiliation strategy used, allowing the YouTube to agree with the overall sentiment of the comment (that the fire is suspicious) but ADJUST the social bonds in the comment to add additional bonds. For example, consider this initiating comment and the replies to it that ADJUST the social bonds in the initiating comment:

83. Main comment: A wave of vandalism, the removal of statues and art, the **slow response** of fire dept, **muslim extremists threatening** Notre Dame, Macron is president. Conclusions made while it's still burning. They **won't look into it.**

    [ideation: vandalism & removal of statues/ attitude: negative VALUATION]

    BAD VANDALISM & REMOVAL OF STATUES BOND

    [ideation: fire department/ attitude: negative CAPACITY]

    LAZY FIRE DEPARTMENT BOND

    [ideation: muslim extremists & Macron/ attitude: negative PROPRIETY]

    EVIL MUSLIMS AND MACRON BOND

    [ideation: authorities/ attitude: negative PROPRIETY]

    Suspicious authorities bond

    Replies:

    i. **Just like** 9/11. Elites **used** Muslims **to carry out the attack.**

    [ideation: 9/11/ attitude: negative VALUATION]

    MANAGE: SUPPORT: WARRANT: ADJUST ↘ SUSPICIOUS FIRE BOND

[ideation: elites/ attitude: negative PROPRIETY]

MANAGE: SUPPORT: WARRANT: ADJUST ⬇️ EVIL ELITES BOND

[ideation: muslims/ attitude: negative CAPACITY]

MANAGE: SUPPORT: WARRANT: ADJUST ⬇️ EVIL MUSLIMS BOND

ii. i know where Michelle Obama was. In Paris sipping champagne while watching the bonfire and peddling her book. **Coincidence or not? .** 🔥

[ideation: Michelle Obama/ attitude: negative PROPRIETY]

MANAGE: SUPPORT: WARRANT: ADJUST ⬇️ EVIL MICHELE OBAMA BOND

iii. **it was the Muslims aided by** traitors in the church.

[ideation: muslims/ attitude: negative PROPRIETY]

MANAGE: SUPPORT: WARRANT: ADJUST ⬇️ EVIL MUSLIMS BOND

[ideation: church traitors/ attitude: negative PROPRIETY]

MANAGE: SUPPORT: WARRANT: ADJUST ⬇️ EVIL CHURCH TRAITORS BOND

iv. I is CLEARLY to **WORK of LUCIFER**, who has The **mooooslums** under his spell.

[ideation: it (fire)/ attitude: negative VALUATION]

MANAGE: SUPPORT: WARRANT: ADJUST ⬇️ SUSPICIOUS FIRE BOND

[ideation: muslims/ attitude: negative PROPRIETY]

MANAGE: SUPPORT: WARRANT: ADJUST ⬇️ EVIL MUSLIMS BOND

The initiating comment (83) lists a number of negative attitudinal targets as evil and suspicious – vandalism, removal of statues, fire department, Muslims and Macron. '*They*' in this case refers to authorities that are supposedly refusing to investigate these negative attitudinal targets. This comment reflects the Anti-Elitist personae, as it predominately delegitimatizes authorities. The replies to this comment share a similar opinion, that the fire is suspicious, but differ in the strength of their attitudinal targets. The first two replies also reflect the Anti-Elitist personae but have ADJUSTED ideational targets. In reply i, the replier agrees that the fire is suspicious, but they add the additional bond of the fire resembling '*9/11*', and elites being the '*evil*' people who caused the fire, with Muslims as their accomplice (in this case, Muslims has negative CAPACITY, as they are being 'used' by elites). In reply ii, Michelle Obama is negative evaluated as she was in Paris when the Notre Dame Fire occurred, and this replier is peddling a

conspiracy theory from this knowledge. Replies iii and iv, represents the White Supremacist personae as explicit negative evaluation is directed towards Muslims. In reply iii, Muslims and church traitors are both negatively evaluated for causing the fire. In reply iv, the fire is evaluated negatively because it is the '*work of Lucifer*'. In addition, Muslims are also negatively evaluated for being under the 'spell' of Lucifer, and for the derogatory spelling of Muslims as '*mooooslums*'. From this example, we can see how ADJUSTING dialogic affiliation works in regards to Conspiratorial comments: the repliers agree with the overall sentiment that the fire is suspicious but differ in their key ideational targets, the people who are blamed for causing the fire.

Repliers that RALLIED around the social bonds in an initiating comment accepted the exact same bond of the initiating commenter. In this sense, RALLYING comments were typically constructed around the same bond – a 'good comment' bond, expressing complete agreement with the initiating commenter. Examples of RALLYING replies are shown here:

84. User 1: **thanks!**

[ideation: comment/ attitude: positive valuation]

MANAGE: SUPPORT: RALLY ↘ GOOD COMMENT BOND

i. User 2: **Quite right .**

[ideation: comment/ attitude: positive valuation]

MANAGE: SUPPORT: RALLY ↘ GOOD COMMENT BOND

ii. User 3: **Good** comment!

[ideation: comment/ attitude: positive valuation]

MANAGE: SUPPORT: RALLY ↘ GOOD COMMENT BOND

iii. User 4: **Extremely**

[ideation: comment/ attitude: positive valuation]

MANAGE: SUPPORT: RALLY ↘ GOOD COMMENT BOND

iv. User 5: **agreed**

[ideation: comment/ attitude: positive valuation]

MANAGE: SUPPORT: RALLY ↘ GOOD COMMENT BOND

v. User 6: **Well** said

[ideation: comment/ attitude: positive valuation]

MANAGE: SUPPORT: RALLY ↘ GOOD COMMENT BOND

vi. User 7: **i feel** you

[ideation: comment/ attitude: positive valuation]

MANAGE: SUPPORT: RALLY ↘️ GOOD COMMENT BOND

All these comments are short and are sometimes completely elide the ideational target. However, it can be inferred that they all express positive attitude towards the initiating comment. Based on the high level of agreement, this means that the repliers are aligning themselves with the same textual personae as the initiating commenter.

Replies that DEFERRED the social bonds in the initiating comment tended to 'laugh off' or 'laugh at' the initiating comment. In other words, 'laughing affiliation' is formed when there is tension and a bond cannot be simply rallied around (Knight, 2010a; Zappavigna, 2018). With the nature of YouTube comments, humour and sarcasm are frequent in comment threads (Thelwall et al., 2012). In the following example, the replier 'laughs at the initiating comment:

85. Initiating comment: Macron **lit** 🔥 **the fire himself**, this way he can blame the yellow jackets 🐝, and save France. (by the way France **is already lost**, it is in the eu)

[ideation: Macron/ attitude: negative PROPRIETY]

MANAGE: SUPPORT: DEFER ↘️ EVIL MACRON BOND

[ideation: France/ attitude: negative CAPACITY]

MANAGE: SUPPORT: DEFER ↘️ LOST FRANCE BOND

[ideation: EU/ attitude: negative PROPRIETY]

MANAGE: SUPPORT: DEFER ↘️ EVIL EU BOND

i. Reply: **lol** like fucking NERO **ha ha!?!**

[ideation: comment/ attitude: positive REACTION]

MANAGE: SUPPORT: DEFER ↘️ ENTERTAINING COMMENT BOND

[ideation: Nero/ attitude: negative PROPRIETY]

MANAGE: SUPPORT: DEFER ↘️ MOCKING NERO AND MACRON BOND

The initiating comment has several main ideational targets. Macron is negatively evaluated (mocked) for supposedly causing the fire and blaming it on the '*yellow jackets*'. France and the EU are also negatively evaluated – France because it has no power, and the EU because it has negatively impacted France. This comment manifests

the Anti-Elitist personae due to these delegitimations of authority. Whether the commenter means this sarcastically or if they are serious, can be difficult to tell without further context, but we can assume that this comment intends to delegitimise authorities based on the negative bonds shared. The reply to this comment, firstly defers the comment by laughing at it, hence forming an ENTERTAINING COMMENT BOND. In the next coupling the reply 'laughs at' the bond that Macron lit the fire, by comparing Macron to the Roman Emperor Nero who was rumoured to have ordered the Great Fire of Rome. This strategy of laughing at the bond and using question marks most aligns with the Sceptic personae that expresses uncertainty. With this example we can see how DEFERRING replies typically treat the initiating comment as a joke, regardless of the initiating commenter's intentions.

DISMISSING dialogic affiliation rejects the initiating commenter's bond and does not propose an alternative bond. There were linguistic similarities with DISMISSING and RALLYING replies, despite the different polarities, these replies were both short and elided ideational targets. DISMISSING replies were typically constructed around two key bonds, the FALSE COMMENT BOND and IGNORANT YOUTUBER BOND. Replies that incorporated the FALSE COMMENT BOND were:

86. User 1: **I don't buy** that .
    [ideation: that (comment)/attitude: negative VALUATION]
    MANAGE: REJECT: DISMISS ↘ FALSE COMMENT BOND
    i. User 2: **not true**
    [ideation: (comment)/ attitude: negative VALUATION]
    MANAGE: REJECT: DISMISS ↘ FALSE COMMENT BOND
    ii. User 3: **BS**
    [ideation: (comment)/ attitude: negative VALUATION]
    MANAGE: REJECT: DISMISS ↘ FALSE COMMENT BOND
    iii. User 4: **Proof?**
    [ideation: (comment)/ attitude: negative VALUATION]
    MANAGE: REJECT: DISMISS ↘ FALSE COMMENT BOND

These comments explicitly state that the comment is false (manifesting the Confrontationist persona) or express scepticism as in 'Proof?' (manifesting the Sceptic persona). The other key bond that DISMISSING replies shared was the IGNORANT YOUTUBER BOND:

87. User 1: **moron**
    [ideation: (YouTuber)/ attitude: negative CAPACITY]
    MANAGE: REJECT: DISMISS ⬇ FALSE COMMENT BOND
    User 2: **Nutcase lol**
    [ideation: (YouTuber)/ attitude: negative CAPACITY]
    MANAGE: REJECT: DISMISS ⬇ FALSE COMMENT BOND
    User 3: **Oh shut up**
    [ideation: (YouTuber)/ attitude: negative CAPACITY]
    MANAGE: REJECT: DISMISS ⬇ FALSE COMMENT BOND

These replies all directly target the YouTuber for their ignorance or insanity, and again manifest the Confrontationist personae. DISMISSING replies were not common with conspiratorial Comments as they were heavily outnumbered by replies that agreed with the initiating comment.

OPPOSING dialogic affiliation disagrees with the initiating comment and provides an alternative coupling. Again, disagreement among conspiratorial initiating comments was not frequent, but this strategy appeared more than the DISMISSING affiliation strategy. The following are examples of OPPOSING social bonds to initiating comments that were of a White Supremacist or Anti-Elitist nature:

88. Replies:
    i. I am a muslim and **didn't Cheer**. Just as many other muslims **have not cheered**. You Guys **are Idiots**
    [ideation: muslims/ attitude: positive PROPRIETY]
    MANAGE: REJECT: OPPOSE ⬇ GOOD MUSLIMS BOND
    [ideation: you guys/ attitude: negative CAPACITY]
    MANAGE: REJECT: OPPOSE ⬇ IGNORANT YOUTUBERS BOND

ii. **not all of us!!!!** I go to church at a Catholic Church. I can assure you <u>our priest and parishioners</u> **don't**. Def can't speak for the Vatican but **don't put that on** <u>all Catholics.</u>
[ideation: us (Catholics)/ attitude: positive PROPRIETY]
MANAGE: REJECT: OPPOSE ⬇️ GOOD CATHOLICS BOND

iii. <u>Thats</u> **bullshit**, are you from Europe, or not, **do not believe everything**....And they have **good** <u>firemen</u>, because ALL firemen are **good WORLDWIDE**, because they help in need and they risk their lives for it.....
[ideation: comment/ attitude: negative VALUATION]
MANAGE: REJECT: OPPOSE ⬇️ BAD COMMENT BOND
[ideation: firemen/ attitude: positive PROPRIETY]
MANAGE: REJECT: OPPOSE ⬇️ GOOD FIREMEN BOND

With reply i, Muslims are positively evaluated, as caring people. The GOOD MUSLIMS BOND opposes the claims in the initiating comment that Muslims cheered when the Notre Dame Cathedral was on fire. In addition, the people commenting White Supremacist statements are negatively evaluated as 'idiots' by the replier. This replier manifests the Educator personae, using additional bonds to explain why a comment is incorrect. In reply ii, we see a similar strategy. The replier opposes the initiating comment that Catholics are associated with the occult. Instead, Catholic are positively evaluated, legitimated by the YouTuber's own person experiences. With reply iii, the replier opposes a initiating commenter that negatively evaluates firefighters. The replier instead negatively evaluates the comment and positively evaluates firemen in terms of social sanction. Thus, in all these examples we can see how opposing replies debunk or call-out what is stated in the initiating comment, presenting more bonds than the dismissing replies. Thus, most of the opposing replies align with the linguistic strategies of the Educator personae.

Lastly, there were replies that did not directly engage with the initiating comment, in the sense that they did not express agreement or disagreement. These replies are labelled as ignoring dialogic affiliation. Examples from the dataset are shown here:

89. User 1: **don't say** huhhhhh like @YouTuber 😂

[ideation: YouTuber/ attitude: negative CAPACITY]

MANAGE: IGNORE ⬇️ IGNORANT YOUTUBER BOND

User 2**: Don't you mean** "…must have been…" or "…must've been …"?

[ideation: you/ attitude: negative CAPACITY]

MANAGE: IGNORE ⬇️ IGNORANT YOUTUBER BOND

User 3: "…must of…" is **grammatically incorrect.**

[ideation: (comment)/ attitude: negative VALUATION]

MANAGE: IGNORE ⬇️ IGNORANT COMMENT BOND

What these examples share in common, is that they do not engage with the YouTuber's bonds but attempt to correct the YouTuber's speaking mannerisms or a commenter's spelling. Again, with these short comments, sometimes ideational targets are not explicitly stated, as with the reply 'grammatically incorrect' that is just attitude with no explicit ideation. These sorts of comments do not align with the personae identified in the initiating comment dataset.

Overall, replies to Conspiratorial comments agreed more frequently with the initiating comment rather than disagreeing with it. ADJUSTING replies were the most common, as YouTuber's preferred to agree with the general gist of the comment but also add their own speculation, by shifting the ideational targets of who is to blame for the fire. In the process of identifying each dialogic affiliation strategy it was also discovered that the repliers typically aligned with the personae identified from the initiating comment dataset. For example, White Supremacists and Anti-Elitists used ADJUSTED or RALLYING replies, Sceptics used DEFERRING replies, and Confrontationists used DISMISSING replies, and Educators used OPPOSING replies. Therefore, this shows some predictability in using dialogic affiliation to identify personae.

7.3.3.2. Replies to Anti-Conspiratorial Comments

Replies to Anti-Conspiratorial comments were less frequent, and hence there was an overall smaller dataset to analyse. These replies more typically disagreed with the

initiating comment, thus there were more instances of opposing and dismissing replies. OPPOSING affiliation was the most common strategy, where repliers disagreed with the initiating comment that attempted to debunk the conspiratorial YouTube video, instead these repliers would add additional bonds of a conspiratorial or xenophobic nature:

90. Replies:

i. **Nothing** in the news is **coincidence**. It's **all messages and spells.**
[ideation: news/ attitude: negative VALUATION]
MANAGE: REJECT: OPPOSE ⬇️ SUSPICIOUS NEWS BOND

ii. **Cheap immigrant workers or Refugees and undercover agents?**
[ideation: immigrants & refugees & undercover agents/ attitude: negative PROPRIETY]
MANAGE: REJECT: OPPOSE ⬇️ SUSPICIOUS IMMIGRANTS/REFUGEES/UNDERCOVER AGENTS BOND

iii. We **don't need any** evidence. Anyone with half a brain will know who **the scum are responsible**
[ideation: evidence/ attitude: negative VALUATION]
MANAGE: REJECT: OPPOSE ⬇️ USELESS EVIDENCE BOND
[ideation: Muslims/ attitude: negative PROPRIETY]
MANAGE: REJECT: OPPOSE ⬇️ EVIL MUSLIMS BOND

With reply i, the replier opposes an initiating comment that states the fire is not suspicious and that coincidences can happen. The replier negatively evaluates the news for its suspicious nature and aligns with the Inciter personae. In reply ii, we also see an additional bond added, the replier here is responding to a initiating comment that discusses workers causing an accidental fire. The replier adds an additional bond of a xenophobic and anti-elitist nature, negatively evaluating immigrant workers, refugees and undercover agents. The question mark in this reply insinuates that they are associated with the fire. In reply iii, the comment directly opposes the initiating commenter's discussion about evidence. This replier delegitimises the importance of evidence by negatively evaluating it, and instead launches into a xenophobic attack of Muslims, that they refer to here as 'THE SCUM'. In all these instances, we see how

opposing replies create additional bonds that align with the Inciter, Anti-Elitist and White Supremacist personae.

DISMISSING strategies were another common strategy among replies to Anti-Conspiratorial comments. In these instances, replies were short and directly called out the initiating commenter:

91. User 1: You are **not a Christian.**
    [ideation: you/ attitude: negative PROPRIETY]
    MANAGE: REJECT: DISMISS ⬇️ FALSE COMMENT BOND
    User 2: **go get help**
    [ideation: (YouTuber)/ attitude: negative CAPACITY]
    MANAGE: REJECT: DISMISS ⬇️ IGNORANT YOUTUBER BOND
    User 3: **Go back to sleep.**
    [ideation: (YouTuber)/ attitude: negative CAPACITY]
    MANAGE: REJECT: DISMISS ⬇️ IGNORANT YOUTUBER BOND
    User 4: **Get back under** your **bridge**

Replies would align with White Supremacist personae, an identity that is concerned with who is righteous, thus by saying someone is not Christian, this would associate them with negative PROPRIETY. Other replies were more aligned with the Inciter personae, who is concerned with people who are not 'AWAKE' to their conspiracy theories, often referring to these people as 'SHEEPLE'.

In comparison to the 'Replies to Conspiratorial Comments', there were less instances of replies that agreed with the initiating comment. Replies that used ADJUSTING strategies, would add additional bonds that provided an additional reason for why the initiating comment was correct in debunking a video. As we see in this instance:

92. Initiating comment: **Ah ah ah** the "mysterious" figure is a fireman 😂😂😂😂 if you play the video for another 5 seconds you can see the rest of the fire brigade joining him. 😂😂😂😂😂😂😂😂😂😂
    [ideation: the video/ attitude: negative VALUATION]

MISLEADING VIDEO BOND

Replies:

i. the head fireman **looked different** because he wore a fire retardant reflective vest over his jacket. **That's all.**

[ideation: footage of firemen/ attitude: positive VALUATION]

MANAGE: SUPPORT: WARRANT: ADJUST ↘️ UNSUSPICIOUS FIRE BOND

ii. Yes, I see **No Robes**, just a hard hat and a hi-vis jacket !

[ideation: footage/ attitude: negative VALUATION]

MANAGE: SUPPORT: WARRANT: ADJUST ↘️ UNSUSPICIOUS FIRE BOND

iii. no link. Just shown to us **live on French TV.** That's all. **No conspiracy**. The internet is full of montages relayed by **not so reliable** medias.

[ideation: live on French TV/ attitude: positive VALUATION]

MANAGE: SUPPORT: WARRANT: ADJUST ↘️ RELIABLE FRENCH TV BOND

[ideation: media/ attitude: negative VERACITY]

MANAGE: SUPPORT: WARRANT: ADJUST ↘️ UNRELIABLE MEDIA BOND

The initiating comment negatively evaluates the YouTube video for being misleading, as the mysterious figure is just a fireman. Similarly, reply i and reply ii agree with the initiating comment but add some additional bonds to explain exactly why it is a fireman in the video. Thus, these comments form an overall UNSUSPICIOUS FIRE BOND, debunking the content of the YouTube videos. In reply iii, the replier also agrees with the initiating comment but provides an additional commentary of why they don't believe the YouTube video, negatively evaluating media that is not reliable, and positive evaluating live French TV. Overall, these replies manifest the Educator personae, as they use additional bonds to explain why the initiating comment is a correct debunk of the YouTube video.

RALLYING strategies directly agreed with what was expressed in the initiating comment, not adding further additional bonds. In similarity to the 'Replies to Conspiratorial Comments' these replies were short and had minimal ideational targets:

93. User 1: **ok**

[ideation: (comment)/ attitude: positive valuation]

MANAGE: SUPPORT: RALLY ↘️ GOOD COMMENT BOND

User 2: (YouTuber) **exactly.**

[ideation: (YouTuber)/ attitude: positive valuation]

MANAGE: SUPPORT: RALLY ↘️ GOOD COMMENT BOND

User 3: **excellent** comment…)

[ideation: comment/ attitude: positive valuation]

MANAGE: SUPPORT: RALLY ↘️ GOOD COMMENT BOND

User 4: **Thanks for that! :)**

[ideation: (comment)/ attitude: positive valuation]

MANAGE: SUPPORT: RALLY ↘️ GOOD COMMENT BOND

User 5: **yeah saw that**. 😊

[ideation: (comment)/ attitude: positive valuation]

MANAGE: SUPPORT: RALLY ↘️ GOOD COMMENT BOND

User 6: thanks. i**nteresting**. it is **good to know** history before making an opinion.

[ideation: (comment)/ attitude: positive valuation]

[ideation: history/ attitude: positive valuation]

MANAGE: SUPPORT: RALLY ↘️ GOOD COMMENT BOND

These replies all featured positive evaluation towards the YouTuber or their comment. As these replies directly align with the initiating comment, this means the repliers are aligned with the same personae as the initiating commenter.

DEFERRING strategies were also evident in the dataset. Some examples include:

94. Initiating comment: OMG another **armchair professional** talking **as if he knows something about firefighting and live streaming**. You should **stop spending so much time on the internet**, and try to **speak with** the people involved in those kind of operations. That might give you **a better understanding. Not worth spending more time** on you.

[ideation: armchair professional/ attitude: negative CAPACITY]

[ideation: you/ attitude: negative CAPACITY]

IGNORANT YOUTUBER BOND

[ideation: people involved in those kinds of operations/ attitude: positive CAPACITY]

CAPABLE AUTHORITIES BOND

i. Reply: **LOL…** the Blue pill or the red pill….

[ideation: comment/ attitude: positive valuation]

MANAGE: SUPPORT: DEFER ↘️ ENTERTAINING COMMENT BOND

95. Initiating comment: Two wrongs **don't make a right!**

[ideation: (video)/ attitude: negative VALUATION]

MANAGE: SUPPORT: DEFER↘️ FALSE VIDEO BOND

i. Reply: Two wongs don't make a white. 😂

[ideation: (comment)/ attitude: positive valuation]

MANAGE: SUPPORT: DEFER ↘️ ENTERTAINING COMMENT BOND

In the first instance (94), the initiating commenter is negatively evaluating and delegitimating the YouTuber for not being an expert in what they've created a video about. Instead, the initiating commenter positively evaluates and legitimises the authorities involved in fighting the fire. The reply to this comment (i) laughs at the bonds shared, instead posing the question whether one wants to remain in ignorance or not. The questioning nature of this reply aligns it with the Sceptic personae. In the second instance (95), the initiating commenter repeats a common proverb, inferring that an allegation of wrongdoing should not be encountered with a similar action. The reply to this comment (i) defers the bond, instead laughing at the comment by rewording right to white, aligning with the white supremacist values of the White Supremacist personae. Thus, these instances of 'laughing at' the bond shared in the initiating comment have added additional meanings to the comment.

Lastly, there were also some instances of IGNORING dialogic affiliation in the dataset, but these were the least common. These instances of ignoring replies asked questions that were unrelated to the initiating comment or made a joke that did not relate to the initiating comment. Some examples are shown here:

96. **Where would the hotwork be done** on a wood and stone building?

[ideation: building/ attitude: negative VALUATION]

MANAGE: IGNORE ⬇️ SUSPICIOUS FIRE BOND

97. **No**, it have been Billy Joel.

[ideation: Billy Joel/ attitude: negative PROPRIETY]

MANAGE: IGNORE ⬇️ SUSPICIOUS FIRE BOND

These examples do not align with the initiating comment that was discussing the Notre Dame Fire and these replies provide irrelevant information, and do not align with the personae discovered in the initiating comment dataset.

Overall, replies to Anti-Conspiratorial comments featured comments that disagreed with the debunking nature of the initiating comments, and resorted to more xenophobic discourse. Thus, OPPOSING dialogic replies manifested the White Supremacist personae, and DISMISSING replies manifested the White Supremacist and Inciter personae. The Educator personae featured amongst ADJUSTING replies, and RALLYING replies were a mixture of Educator and Confrontationist personae. DEFERRING replies were also a mixture of either the White Supremacist or Sceptic personae. In all, the replies to Anti-Conspiratorial comments featured more disagreement and xenophobic discourse.

7.3.2.2. Summary of Results

The analysis of replies to Conspiratorial and Anti-Conspiratorial comments has revealed how the personae from the initiating comments dataset can be mapped according to dialogic affiliation strategies, and the dominance of interaction with conspiratorial discourse by textual personae in the dataset. The main findings include that the Educator personae adopts OPPOSING or ADJUSTING dialogic affiliation strategies as they need to table additional bonds in order to educate their audience, whilst the Confrontationist personae adopts the RALLYING or DISMISSING dialogic affiliation strategy in order to censure the YouTuber. Other personae such as the Xenophobe, Conspiracist

and Anti-Media were found across all dialogic affiliation strategies rather than confined to a specific dialogic affiliation strategy. The Sceptic personae aligned with the DEFERRING, ADJUSTING or OPPOSING affiliation strategies. In *Table 7-6* an overview of the dataset is provided. As previously discussed, Conspiratorial comments featured more agreeing dialogic affiliation strategies than anti-conspiratorial comments, and an overall greater number of replies. The analysis and numbers in this table indicate that there is not much interaction between textual personae aligning with conspiratorial views and textual personae who want to debunk these videos and combat xenophobia.

Type of Comment	**Number of Initiating Comments with Replies**	**Number of Replies (Total)**	**Macro-Level Affiliation Strategies**	
Conspiratorial	135	705	Agree	572
			Disagree	22
			Ignore	111
Anti-Conspiratorial	89	249	Agree	85
			Disagree	150
			Ignore	14

Table 7-6 - Replies Analysis Dataset

7.4. Comparing Personae

This section will compare personae across the Momo Challenge and Notre Dame Fire case studies. Analysing the patterning of affiliation and legitimation strategies derived from these case studies is important in order to provide a comprehensive understanding how this method can be applied to information disorders in different contexts and addresses the final research question of this thesis. This section will also highlight additional research that complements the findings derived from both case studies.

7.4.1. Political versus Non-Political Personae

As previously mentioned, the two case studies feature instances of two different discourses: non-politically motived (the Momo Challenge) and politically motivated (the Notre Dame Fire). In addition, whilst these case studies consist of different macro-genres of information disorder in terms of a non-politically motivated internet hoax (Wardle and Derakhshan, 2017) versus politically motivated conspiracies, there are similar linguistic patterns that can be identified from these case studies. Namely, these linguistic patterns refer to dominant affiliation strategies and ways of evaluating. In terms of differences, each case study has unique targets of evaluation, thus resulting in the different personae categories that emerge. Nonetheless, as this section will show, the personae categories identified from both case studies can be grouped together to reveal interesting insights into how communities engage with misinformation and disinformation.

7.4.1.1. Sceptics

The Sceptic persona featured in both the Momo Challenge and Notre Dame Fire case studies. The key linguistic feature of this persona across the case studies was the use of EMBELLISHING AFFILIATION – expressing an open-minded or uncertain point of view rather than expressing a definite stance. The Sceptic persona questions a statement rather than doubting knowledge altogether. In other words, sceptics are inquirers. Similarly, the majority of Sceptic persona comments feature questions that open the scope of possibilities rather than delimiting them. However, within these comments there are also evaluations that at times pre-empt the answer to the question or highlight the stance that the commenter is positioned towards.

In these two examples, the Sceptic persona is realised via the use of EMBELLISHING AFFILIATION:

98. I would love to **investigate** this, does anyone know where I can find Momo's number or official videos? Have they all been taken down?

[ideation: this/ attitude: invoked negative VALUATION]

Convoke: Marshal (anyone) ↘️ SUPERIOR KNOWLEDGE BOND

Finesse: Embellish (would, ?) ↘️ SAFE MOMO BOND

Authorisation: Impersonal: Technological

99. I'm **not sure what I'm seeing** to make a clear judgment.

[ideation: video/ attitude: invoked negative VALUATION]

Finesse: Embellish (not sure) ↘️ UNCLEAR VIDEO BOND

De-Authorisation: Impersonal: Technological

The Example 98, taken from the Momo Challenge case study, uses EMBELLISHING AFFILIATION to open possibilities through expanding engagement strategies such as "*I would love*", "*does anyone know*" and the repetition of question marks. Rather than making any direct assertions, questioning is used to trigger further inquiry. There is also invoked NEGATIVE VALUATION. By questioning the Momo challenge and asserting the difficulty in finding Momo's number or official videos, this negatively evaluates the veracity of the Momo challenge. Similarly, in example 99, taken from the Notre Dame Fire case study, EMBELLISHING AFFILIATION is evoked via the expanding engagement strategy of "*not sure*". This opens uncertainty regarding the veracity of the Notre Dame Fire video. Also similar to the previous example, there is an INVOKED NEGATIVE EVALUATION of the YouTube video by questioning its veracity. These examples show how EMBELLISHING AFFILIATION is a key feature of the Sceptic persona, and how these strategies exist in two completely different case studies.

Beyond these two case studies, the Sceptic persona has also been identified in other case studies. For example, in analysing comments on Russia Today (RT) videos about the Skripal poisoning (Inwood and Zappavigna, forthcoming), the Sceptic persona was distinguished by its use of EMBELLISHING AFFILIATION:

100. So this was a Russian lab that **found the bz?**

[ideation: Russian lab/ attitude: invoked negative VALUATION]

Finesse: Embellish (?) ↘️ UNTRUSTWORTHY RUSSIAN LAB BOND

In this example, the use of the question mark (?) EMBELLISHES the ideational target (Russian lab), thus expanding the range of possible meanings by posing a question rather than hindering meanings via a definitive statement. However, in understanding the broader contextual meanings surrounding the statement, the Russian lab has an INVOKED NEGATIVE ATTITUDE (i.e. Russian labs are evaluated as untrustworthy). In a similar pattern to the Momo Challenge and Notre Dame Fire case studies, Sceptics used more negative invoked evaluations.

The Sceptic persona was identified in both political and non-political discourse with similar linguistic strategies (EMBELLISHING AFFILIATION and invoked negative evaluations). Whilst it appears at first that the Sceptic persona expresses an open-minded stance, a close analysis of Sceptic persona comments reveals invoked negative evaluations. This personae category was one of the most consistent in terms of linguistic strategies across political and non-political discourse (as identified in Table 7-1 and Table 7-4 in this chapter).

7.4.1.2. Confrontationists and Critics

The Confrontationist and Critic personae share in common a direct critique of fellow commenters and videos creators via negative evaluations. Confrontationists in the Notre Dame Fire case study call out someone as stupid or false without explaining further detail. On the other hand, Critics in the Momo Challenge case study directly criticise the media, established institutions and older generations, but these responses range from simply 'calling out' the target to explanations with further details and evaluations.

Both the Confrontationist and Critic personae were ultimately sharing truthful information, as the Confrontationist critiqued the false information of YouTubers and

commenters, and the Critic critiqued the scare-mongering behaviour of institutions and people that propagated the false assertion that Momo is real. For example, in this comment from the Momo Challenge dataset, the news media is criticised for overreacting to the Momo Challenge and misunderstanding internet culture:

101. Warning⚠️The news media just discovered Momo. Mass hysteria **overreaction** to **misunderstood internet culture** imminent. Prepare for **wide-spread moral panic, demonetization** video takedowns.

[ideation: news media/ attitude: invoked negative CAPACITY]

[ideation: (media)/ attitude: negative CAPACITY]

[ideation: moral panic / attitude: negative VALUATION]

Temper: Modulate (⚠️, just) ↘️ FAKE MEDIA BOND

Moral Evaluation: Evaluation

De-Authorisation: Impersonal: Technological

This comment calls out the assumptions made in the YouTube video (that the Momo Challenge is dangerous) by negatively evaluating the "*news media*" for their "*mass hysteria overreaction*". This particular negative evaluation is also a strategy of delegitimising because it is targeted towards a particular institution and aims to discredit its validity. In another example from the Notre Dame Fire dataset, a similar strategy of calling out the YouTuber's video is adopted:

102.You guys are **full of shit**, the narrative you're pushing **makes absolutely no sense.**

[ideation: guys/ attitude: negative VERACITY]

[ideation: narrative/ attitude: negative VALUATION]

CONVOKE: DESIGNATE (you guys) ↘️ LYING YOUTUBER BOND

TEMPER: MODULATE (absolutely) ↘️ BAD NARRATIVE BOND

De-Authorisation: Personal

De-Mythopoesis: Single Determination

In this example, "*guys*" (the creators of the YouTube video) are evaluated negatively for spreading false narratives. The CONVOKING AFFILIATION of "*you guys*" enhances the sense of calling out their behaviour, alongside the TEMPERING AFFILIATION of "*absolutely*". Similar to the first example, this calling out is also a delegitimization strategy,

delegitimising both the person (*you guys*) and the narrative (*makes absolutely no sense*). Therefore, these two comments share in common a calling out of a person or institution, in other words, negatively evaluating a specific target for spreading false information. In particular, it is the specific evaluative strategy of NEGATIVE VERACITY or NEGATIVE VALUATION, that needs to be attached to the ideational target, forming lying or fake bonds, that distinguishes the Confrontationist and Critic personae to other personae that have negative ideational targets but instead focus their attitude on NEGATIVE PROPRIETY, that is, morals rather than questions of veracity.

In further research, the textual personae of the "Heckler" has been identified. This was based on an analysis of comments on Russia Today (RT) videos of the Skripal poisoning (Inwood and Zappavigna, forthcoming). Similar to the Confrontationist and Critic, the Heckler negatively evaluates a target, in this case the Russian people:

103. RUSSIA NICKNAME IS THE **PAPER BEAR**

[ideation: Russia/ attitude: negative CAPACITY]

TEMPER: MODULATE (PAPER) ↘ POWERLESS RUSSIAN BOND

CONVOKE: DESIGNATE (RUSSIA) ↘ POWERLESS RUSSIAN BOND

In this example, Russia is negatively evaluated, as the "*paper bear*" has connotations with a weak state that pretends to be strong. The commenter is heckling (aggressively targeting) the creator of the video Russia Today (RT). The TEMPERING AFFILIATION (capital letters) strengthens the statement whilst the CONVOKING AFFILIATION (naming Russia) directly identifies the target. Thus, the Heckler personae shares a similar calling out strategy as the Confrontationist, but also embodies CONVOKING AFFILIATION to directly name the target.

Overall, Confrontationists and Critics both focus on negative evaluations that directly call out the target for spreading false information. Whilst Confrontationists provided negative evaluations with little explanation, Critics expressed negative evaluations with further details. In considering the case studies mentioned, there was in each case study textual personae that negatively evaluated other commenters or video creators for

spreading false information or making disturbing statements. Thus, based on this analysis, a pattern emerged of personae in comment threads that always call out the dominant statements made in the videos they are responding to.

7.4.1.3. Educators and Connoisseurs

Educators and Connoisseurs both enacted a pedagogic role in discourse. In the Momo Challenge case study, Connoisseurs provided an insider commentary on internet culture, thus educating readers that Momo is just a sculpture and nothing to fear. In the Notre Dame Fire case study, Educators focused on explaining why the conspiratorial Notre Dame Fire videos were false and provided multiple evaluations, including negative evaluations of the veracity of information provided and positive evaluations of people as inherently good.

Educators and Connoisseurs have a similar function in portraying themselves as experts in the subject discussed or being able to refer readers to the experts in the matter discussed. Etymologically to educate means to 'draw out' or 'unfold the powers of the mind' (Harper, n.d.), and in a similar sense Educators and Connoisseurs express multiple bonds (rather than single evaluations) in order to explain an issue in detail. Similarly, a Connoisseur refers to a 'critical judge of any art', an 'expert', or in other words, someone who is 'well versed in a subject' (Harper, n.d.). In the Momo Challenge case study, Connoisseurs were well versed in internet culture and used their comments to explain to readers the subtle meanings behind the Momo Challenge. For example, in the following:

104. The momo thing is **fake**. If you go onto know your meme, it explains how it's **a sculpture on a stand'n shit.**

[ideation: momo/ attitude: negative VALUATION]

[ideation: it (momo)/ attitude: invoked positive APPRECIATION]

Finesse: Embellish (it explains) ↘ FAKE MOMO BOND

Temper: Modulate ('n shit) ↘ FAKE MOMO BOND

Convoke: Marshal (you) ↘ SUPERIOR KNOWLEDGE BOND

Authorisation: Impersonal: Technological

This comment negatively evaluates Momo as false. Whilst the Critic persona was also concerned with questions of negative veracity, what distinguishes the Connoisseur persona is the explanation function of the comment. Rather than simply calling out Momo as false, the Connoisseur persona features further bonds such as the "SUPERIOR KNOWLEDGE BOND", providing further information about why Momo is fake, rather than just calling out Momo as fake. The Connoisseur persona also used further legitimation strategies to validate their claims, for example, how this comment states that one can go onto "*know your meme*" in order to ascertain the Momo is fake.

A similar pedagogic function exists with the Educator persona in the Notre Dame Fire dataset. This persona was also concerned with questions of veracity but then added additional evaluations in order to educate the reader rather than simply calling out something as false:

105. Pretty sure that "allahu ackbar" audio was **dubbed over fire footage.** It's practically a **meme**, dubbing "allabu ackbar" over any disaster and "fail".

[ideation: Allahu ackbar audio/ attitude: invoked negative VALUATION]

[ideation: it (Allahu ackbar audio)/ attitude: invoked negative VALUATION]

FINESSE: DISTIL (pretty sure, it's) ↘️ MISLEADING VIDEO BOND

De-Authorisation: Impersonal: Technological

In this example, the "*Allahu ackbar audio*" is negatively evaluated as fake. However, this comment goes one step further by explaining why this negative evaluation occurs. Legitimation strategies are also essential in this comment, as the audio is delegitimised because it is a meme. Thus again, rather than simply calling out the audio, further bonds and legitimation strategies seek to educate the reader on the decision behind this negative evaluation.

A similar textual persona to the Educator and Connoisseur is the Anti-Propagandist. The Anti-Propagandist was identified the comment section of RT (Russia Today) videos

on the Skripal poisoning (Inwood and Zappavigna, forthcoming). For example, in the following comment RT as a propaganda channel is directly called out:

106. LOL, exclusive interview – RT is a **Russian owned and run propaganda channel** and Peskov **is their troll.**

[ideation: RT/ attitude: negative VERACITY]

FALSE RT BOND

[ideation: Peskov/ attitude: negative VERACITY]

Temper: Modulate (LOL) ↘ LYING PESKOV BOND

Finesse: Distil (is) ↘ LYING PESKOV BOND

This comment evaluates RT with NEGATIVE VERACITY for being a '*Russian owned and run propaganda channel*', Peskov, a Russian diplomat, is evaluated with NEGATIVE VERACITY for being a '*troll*'. The TEMPERING AFFILIATION of '*LOL*' laughs off the idea of an exclusive interview, instead evaluating RT as false information. DISTILLING AFFILIATION also has a role in this comment, as the '*is*' makes these claims definitive without room for negotiation. From this example it can be observed how the Anti-Propagandist personae in similarity to the Educator and Connoisseur is focused on evaluations of veracity, as a way of informing readers about false information.

The Educator personae was also examined in comments on videos promoting COVID-19 conspiracies (Inwood and Zappavigna, 2022a). Similarly, this Educator persona debunked content in the videos. Comments by Educators for example explicitly indicated that the video was false:

107. This video is **propaganda** and a push for a libertarian ideology. Come to NYC and start complaining about your **previous** civil liberties....this is a global pandemic and thousands are dying it's **horrible**. The shutdown is to **reduce the curve** so we don't all die dude **Wake up**, this is **real**.

[ideation: YouTuber's creation of video/ attitude: inscribed negative T-VERACITY]

Finesse: Distil (is) ↘ FAKE VIDEO BOND

[ideation: complaining about civil liberties/ attitude: negative VALUATION]

Convoke: Marshal (your) ↘ IGNORANT PERSON BOND

[ideation: pandemic/ attitude: negative REACTION]

Finesse: Distil (it's) ↘ DANGEROUS VIRUS BOND

[ideation: shutdown/ attitude: positive VALUATION]

Temper: Modulate (all) ↘ REAL VIRUS BOND

[ideation dude/ attitude: negative CAPACITY]

Convoke: Marshal (wake up) ↘ IGNORANT PERSON BOND

[ideation: this (experiencing a pandemic)/ attitude: invoked positive T-VERACITY]

Finesse: Distil (is) ↘ REAL VIRUS BOND

In this comment, the YouTuber's creation of the video is associated with NEGATIVE T-VERACITY because it is overtly referred to as '*propaganda*', therefore a FAKE BOND is shared via DISTILLING AFFILIATION limiting other possibilities. The user associates the pandemic with NEGATIVE REACTION and a DANGEROUS VIRUS BOND. '*Dude wake up*' infers an IGNORANT BOND via MARSHALLING AFFILIATION as the YouTuber doesn't understand the severity of the pandemic, and 'DUDE' here has NEGATIVE CAPACITY. Lastly, the experience of living in a pandemic is coded with POSITIVE T-VERACITY and a REAL BOND via DISTILLING AFFILIATION is shared as the user highlights the importance of the shutdown and again limits false interpretations.

Educators and Connoisseurs are both characterised by making multiple evaluations and using DISTILLING AFFILIATION strategies in order to definitively alert readers to the dangers of false information. Whilst Educators more broadly discussed the falsity of the information provided in the videos they were responding to, Connoisseurs expressed more specific knowledge by highlighting their knowledge of internet culture. Thus, these personae differed in legitimation strategies, however, both used DISTILLING AFFILIATION in order to make definitive claims.

7.4.1.4. Nationalists

From the array of textual personae uncovered, Nationalists were an outlier. Nationalists were uncovered in the Momo Challenge dataset but not in the Notre Dame Fire dataset. In other words, there were textual personae in the Momo Challenge dataset who identified themselves according to their nationality, however, in the Notre Dame Fire dataset this was not as clear. Instead, in the Notre Dame Fire dataset commenters

focused on establishing a dichotomy of white supremacy versus threatening immigrant groups.

In the Momo Challenge dataset, Nationalist textual personae are distinguished by their use of CONVOKING AFFILIATION, that is, the use of vocatives or designation to call out to their community or name their community. In this dataset, Nationalism is portrayed as positive, as commenters name their nationality in order to ascertain that Momo is nothing to fear and create a sense of unity. For example, in the following comment the commenter aligns themselves with the Philippines:

108. Here in Philippines momo is just a **meme** for us
[ideation: momo/ attitude: invoked positive AFFECT]
Convoke: Designate (Philippines, us) ↘ SUPERIOR NATION BOND
Temper: Modulate (just) ↘ SAFE MOMO BOND
De-Authorisation: Impersonal: Technological
Authorisation: Personal

"*Here in Philippines*" has a DESIGNATING AFFILIATION function, aligning the commenter with a particular community. This alignment also evokes a sense of nationalism, as the commenter is tempering down the severity of Momo, by stating Momo is "*just a meme*", this assumes a SAFE MOMO BOND, rather than a DANGEROUS MOMO BOND that Western countries like the USA tend to adopt. Therefore, a SUPERIOR NATION BOND arises from this invoked evaluation of Momo and knowledge of how Momo is interpreted in different cultural contexts.

The alignment of social bonds to a nationality in the Momo Challenge dataset appears quite different to how social bonds that evoke nationalist tendencies are expressed in a more political context. For example, in comments on RT videos about the Skripal poisoning (Inwood and Zappavigna, forthcoming), Nationalism is used to threaten another group rather than create unity:

109. Russia **will make** the west to **speak the truth**
[ideation: Russia/ attitude: positive CAPACITY]

POWERFUL RUSSIA BOND
[ideation: the West/ attitude: negative VERACITY]
Finesse: Distil (will make) ↘️ LYING WEST BOND

In this comment, Russia is portrayed with POSITIVE CAPACITY because it '*will make*' the West to speak the truth. '*Will make*' is an example of DISTILLING AFFILIATION as it is definitive, leaving little room for negotiation. In contrast to the power and truthfulness of Russia, the West is portrayed with NEGATIVE VERACITY, for refusing to speak the truth. These types of Nationalist comments emphasise a powerful and righteous Russia in comparison to a deceiving West.

Overall, the Nationalist personae appears to differ in non-political versus political contexts. Similarly, to the competing definitions of nationalism previously explored, nationalism can be thought of as positive community-building in order to erase fears (in the case of the Momo Challenge case study) or as a way to criticise and intimate other communities (in the case of the RT video dataset). The Nationalist personae was absent in the Notre Dame Fire case study as commenters did not align themselves with particular nations, rather alignment tended towards particular religions and races.

7.4.1.5. Moralisers, Anti-Elitists, Anti-Media, and White Supremacists

A group of textual personae identified across the Momo Challenge and Notre Dame Fire case studies focused on negatively evaluating particular targets and using TEMPERING AFFILIATION to emphasise claims. In the Momo Challenge case study, this refers to the Moralisers who direct negative judgement towards parents and hackers and platforms such as YouTube. In the Momo Challenge case study Moralisers solely use negative judgement, particularly NEGATIVE CAPACITY and NEGATIVE PROPRIETY, to criticise parents for being lazy, hackers for being evil, and platforms for being morally inept. For example, in the following, negative judgement is directed towards parents:

110. Oh no! now all u **lazy** parent can't use YouTube to entertain ur kids. Now it **gna have to get involved with ur kid life's Instead of leaving them in the other room on the iPad. God bless** u momo

[ideation: parents/ attitude: negative TENACITY]

[ideation: (it) parents/ attitude: negative PROPRIETY]

[ideation: momo/ attitude: invoked positive JUDGEMENT]

Convoke: Marshal (all u lazy parent) ↘ OUT GROUP COMMUNITY OF BAD PARENTS

Finesse: Embellish (have to get involved, instead of) ↘ BAD PARENT BOND

Moral evaluation: Evaluation

De-Authorisation: Role Model

The ideational target of parents is linked to both NEGATIVE TENACITY (for being lazy) and NEGATIVE PROPRIETY (for their immoral behaviour in not being involved in their kids' lives). The repetition of negative evaluations towards parents insinuates this comment as an act of hatred towards lazy parents as well as that of moral judgement.

In the Notre Dame Fire case study, the macro-personae of hate speech defines Anti-Elitists, Anti-Media and White Supremacists who negatively target elites, mainstream media, or immigrants. In the following example from the Notre Dame Fire dataset, Muslims are the target of negative evaluations:

111. Muslims **sure were in celebration** during the blaze. There's several videos circulating online that show them **screaming allahu akbar** at the fire. **They laugh about it on Facebook.** They **did it absolutely.**

[ideation: Muslims & they/ attitude: negative PROPRIETY]

[ideation: Muslim immigrant/ attitude: negative PROPRIETY]

Temper: Modulate (sure, several, absolutely) ↘ EVIL MUSLIMS BOND

Finesse: Distil (sure were, did it) ↘ EVIL MUSLIMS BOND

Authorisation: Impersonal: Technological

Similar to the Momo Challenge dataset, these negative evaluations are in the realm of hate speech and moralisation. Muslims are negatively targeted for supposedly laughing about the Notre Dame Cathedral being on fire and for supposedly causing the fire. These

repeated negative evaluations of a particular ideational target again situates this comment in the range of hate speech.

In other research, another textual persona that aligns with the personae previously mentioned is the Anti-Globalist (Inwood and Zappavigna, 2022a). The Anti-Globalist employs negative attitudinal resources targeted at those who supposedly hold a high degree of power: elites, the deep state, and globalist organisations like the UN and WHO. The elites were often blamed for creating problematic vaccines:

112. Hey **elites!** Guess what? You can take your ID and vaccine and **shove it up your arse!**
[ideation: elites/ attitude: invoked negative PROPRIETY]
[ideation: ID and vaccine/ attitude: negative REACTION]
Convoke: Marshal (hey elites!) ↘️ EVIL ELITES BOND

In this example, a CONVOKING strategy is directed towards elites (‘*hey elites*’) and elites are associated with NEGATIVE PROPRIETY for being a common anti-bonding icon. The ‘*ID and vaccine*’ has NEGATIVE REACTION. An EVIL ELITES BOND is shared here, as elites have created the negatively evaluated ‘*ID and vaccine*’.

Overall, these categories of personae can be defined by negative judgement towards specific targets. In most cases, TEMPERING AFFILIATION was also used to emphasise claims, however, there are also instances of CONVOKING AFFILIATION addressing the particular groups that the negative judgement is directed towards. These categories of personae represent hate speech, thus, these personae are sharing misinformation, on a spectrum from problematic content to false information.

7.4.1.6. Myth Spreaders and Inciters

Myth Spreaders and Inciters spread false information in a narrative format. These types of personae typically use monoglossic language (DISTILLING AFFILIATION), TEMPERING AFFILIATION to emphasise claims, and inconsistent statements not following any pattern of logic. In the Momo Challenge case study with the Myth Spreader persona, ‘Momo’

can be likened to a supernatural being that people spread rumours about, repeating what they have overheard, typically at school or via a parent, as their supposed evidence, as in the following comment:

113. Apparently the principle in my school caught two kids cutting themselves with pencils in the bathroom because of the Momo challenge, and sent an email warning all parents about Momo. This is some **scary** shit.
[ideation: Momo challenge/ attitude: invoked negative JUDGEMENT]
[ideation: this shit/ attitude: negative REACTION]
Finesse: Embellish (apparently) ↘ DANGEROUS MOMO BOND
Temper: Foster (warning all) ↘ DANGEROUS MOMO BOND
Mythopoesis: Cautionary Tale

In this comment, it has a narrative function by recounting what was supposedly overheard by the commenter. The MYTHOPOESIS legitimation strategy of this comment situates it as a cautionary tale, a warning of the dangers of the Momo challenge. TEMPERING AFFILIATION strengthens the impact of the story as well. Thus, we can see the role of storytelling in the creation of the Myth Spreader persona.

In the Notre Dame Fire case study, narratives are used to spread false information. Inciters create conspiracies about how the Notre Dame Fire started. These conspiracies involved a SUSPICIOUS FIRE BOND leaving the narrative with a sense of mystery regarding the exact cause of the fire.

114. The **absurdly premature** denial of arson is **absolute proof** that the Deep State **set fire to Notre Dame.**
[ideation: denial of arson/ attitude: invoked negative VERACITY]
[ideation: Deep State/ attitude: negative PROPRIETY]
Temper: Modulate (absurdly premature) ↘ LYING AUTHORITIES BOND
De-Authorisation: Commendation: Expert
Mythopoesis: Single Determination
Temper: Modulate (absolute) ↘ EVIL DEEP STATE BOND

In occurrence with the Momo Challenge comment example, the MYTHOPOESIS LEGITIMATION strategy, to recount the SINGLE DETERMINATION made by the Deep State (a well know ideational target of conspiracists) situates this comment as relying on storytelling and narrative formats in order to project its message. TEMPERING AFFILIATION (*absolute proof*) also strengthens the impact of this story. Thus, similar strategies of storytelling and TEMPERING AFFILIATION appear as a pattern in the comments of both Myth Spreaders and Inciters.

A Conspiracist personae was also evident in YouTube comments about the Skripal poisoning (Inwood and Zappavigna, forthcoming). Take for example the following:

115. This was obviously a **staged hoax** deliberately designed to **demonise** Russia
[ideation: this/ attitude: negative VALUATION]
[ideation: Russia/ attitude: positive PROPRIETY]
Finesse: Distil (obviously and deliberately) ↘ FALSE EVENT BOND & INNOCENT RUSSIA BOND

In this example, the Skripal poisoning (*this*) is negatively evaluated as a '*staged hoax*', thus forming a FALSE EVENT BOND. The finessing of '*obviously*' strengthens the impact of this FALSE EVENT BOND. In the next part, Russia is treated with positive PROPRIETY because it has been demonised (with the comment implicitly inferring the West is demonising Russia). Again, the FINESSING affiliation of '*deliberately*' strengthens the impact of this statement, and makes it appear as a definitive position, without negotiation. The Conspiracist persona is distinguished by this contrasting patterning of a FALSE EVENT BOND (that evaluates a thing rather than a person/country) and an INNOCENT RUSSIA BOND, with DISTILLING affiliation making these bonds appear definitive.

Overall, Myth Spreaders and Inciters spread conspiratorial content, that is, false information or speculation in a narrative format. In the Momo Challenge, this manifested as myths or hearsay, in the Notre Dame Fire, this manifested as inciting suspicion about the fire via conspiratorial narratives. In these case studies, suspicious bonds and negative evaluations, with TEMPERING and DISTILLING AFFILIATION, were common for these sorts of personae.

7.4.2. Overview of Political Versus Non-Political Personae

Overall, in comparing textual personae from non-politically motivated and politically motivated case studies, several key linguistic patterns emerged that have been termed as "macro-categories". Six macro-categories were identified after compiling the analytical results of the two case studies. These macro-categories were identified from the similar evaluation strategies and textual telos (the pursuit of goals) shared by textual personae (Martin and White, 2005: 49). For example, Critics and Confrontationists both negatively evaluated individuals in order to establish the veracity of claims in a YouTube video. Whilst these personae were labelled differently in order to reflect the ideational targets of the specific case study explored, Critics and Confrontationists both share the same goal in censuring false information.

A case study does not need to reflect all six macro-categories identified. For example, Nationalists were identified in the Momo Challenge case study only, which initially appears surprising as Nationalism is often associated with political discourse. However, the particular political case study selected, the Notre Dame Fire, did not contain commenters that made their particular nationality evident. Rather, commenters in the Notre Dame Fire dataset targeted non-Western cultures and individuals through hate speech that did not make particular reference to commenters' own nationality. The Nationalist personae was also a point of difference among the case studies for the reason that in the Momo Challenge the Nationalist personae spread truthful information about the Momo Challenge, in comparison other studies have focused on how Nationalism can be used to spread false information (Busby, 2017; Fuchs, 2019; Green, 2021).

The Notre Dame Fire case study featured a greater range of types of personae sharing false information (Anti-Elitists, Anti-Media, White Supremacists, Inciters) compared to the Momo Challenge case study (Myth Spreaders and Moralisers). This is because the Notre Dame Fire dataset featured a broader range of negatively evaluated ideational targets (i.e., the targets of hate speech and different political ideologies engaged with) compared to the Momo Challenge, that perhaps due to its unpolitical nature, had fewer

ideational targets. Thus, these are key points of difference among non-political versus political discourse.

Lastly, the macro-categories identified should not be seen as an exhaustive list of the persona categories that may occur across information disorder discourse. New categories may emerge depending on the case study. However, this work does provide a method for categorising the patterns that emerge in the linguistically enacted communicative strategies and evaluations in YouTube comment sections. Table 7-7 provides a summary of the key features discussed in Section 7.4. The 'Macro-Category' column provides a non-technical definition of the communicative patterns unique to each macro-category. The 'Momo Challenge Case Study' and 'Notre Dame Fire Case Study' columns on the other hand demonstrate the key ideational targets and predominant affiliation strategies unique to each persona in the specific case study.

Macro-Category	Momo Challenge Case Study	Notre Dame Fire Case Study
Scepticism Questioning the veracity of the YouTube video	**Sceptics** EMBELLISHING AFFILIATION, invoked negative evaluations	**Sceptics** EMBELLISHING AFFILIATION, invoked negative evaluations
Criticism Criticising the veracity of a YouTube video or individual	**Critics** Negative attitude towards the media, established institutions, and older generations that falsely claimed the Momo Challenge is real	**Confrontationists** Negative attitude towards (calling out as stupid/false) a YouTuber or YouTube video
Education and Expertise Educating and providing expertise to an audience	**Connoisseurs** Providing insider commentary on internet culture, FINESSING AFFILIATION to ascertain that Momo is an object not a person, multiple social bonds	**Educators** Educating readers about why a video is false, multiple social bonds, more positive evaluation to counteract hate speech
Nationalism Emphasising the importance of geographic location/nationality in determining the veracity of information	**Nationalists** CONVOKING AFFILIATION to state one's country has a superior understanding of Momo	Not evident in dataset – commenters did not make obvious their nationality
Hate Speech Negative judgement towards a person or group of people encouraging hate	**Moralisers** Negatively judging parents and hackers, and YouTube	**Anti-Elitists, Anti-Media, White Supremacists** Negatively evaluating authorities, mainstream media or non-Western cultures
Storytelling and Conspiracy Spreading false information in a narrative format	**Myth Spreaders** Inconsistent stories about the Momo Challenge, TEMPERING AFFILIATION to emphasise claims	**Inciters** Inconsistent stories about how the fire started, suspicious fire bond and unnamed ideational targets

Table 7-7 - Overview of Personae Comparison

7.3. Conclusion: Relevance and Other Applications of Textual Personae

This chapter has constructed and applied a method for identifying textual personae in YouTube comment threads that engage with information disorders. By undertaking an analysis of affiliation and legitimation patterns, various textual personae in the two case study datasets were identified. These personae displayed recurrent patterns in evaluative meanings, affiliation strategies, and legitimation strategies. Additionally, as shown via the bond cluster diagrams, textual personae could be grouped according to positive versus negative evaluations and the absence of shared bonds resulted in animosity among particular personae. By analysing the linguistic patterns unique to textual personae identified in the Momo Challenge and Notre Dame Fire case studies, macro-categories of textual personae were also developed. These macro-categories provided a holistic overview of the function of textual personae in comment sections, and these findings were additionally supplemented with the results from pilot studies. Whilst the macro-categories do not provide an exhaustive list of every type of textual personae that could appear in YouTube comment sections, the method discussed provides a means for understanding and categorising in SFL terms how textual personae are linguistically enacted across large datasets.

The implications of this research beyond the particular case studies explored in this chapter is beneficial to qualitative and quantitative studies regarding discourses of information disorder. By using linguistic analysis to identify textual personae in discourse, this means that targeted solutions can be developed for specific communities, addressing particular social bonds, rather than applying a one-size-fits-all approach to understanding how information disorders unfold in online communities. As a bond cluster diagram illustrates via ying-yang symbols representing the coupling of ideational and attitudinal meaning, information disorder discourses can be related to a range of different social bonds. It is important to delve further into the linguistic construction of these social bonds and how different social bonds relate to each other in order to understand these complex online interactions and the core values at stake. In addition, understanding the textual personae in comment sections can aid

in the creation of manual and automated detection processes for identifying the specific networks of values that charge information disorders. SFL approaches to studying information disorders can also be useful in designing educational materials so that students can linguistically identify the language used by those engaging with propaganda and prevent themselves from becoming radicalised (Szenes, 2021).

The method for identifying textual personae detailed in this chapter has previously been applied to the genre of internet hoaxes (Inwood and Zappavigna, 2021), conspiratorial and white supremacist discourse (Inwood and Zappavigna, 2023), Russian propaganda (Inwood and Zappavigna, forthcoming), and COVID-19 conspiracy theories (Inwood and Zappavigna, 2022a). Whilst some personae such as 'Sceptics' are shared across these different case studies, other personae categories are different, particularly in comparing non-political versus political discourse. Future research might apply this method to more case studies, encompassing a range of different issues (as will be further discussed in *The Conclusion Chapter*). The method could also be refined to identify the patterns more closely in language most relevant to these potential domains, and further specify the relationships between affiliation strategies and bonds.

8. CHAPTER EIGHT CONCLUSION: A MULTI-LAYERED INTEPERSONAL APPROACH TO STUDYING INFORMATION DISORDERS

8.1. Introduction

This chapter will summarise the key findings of this thesis in terms of how a multimodal affiliation and legitimation analysis reveals the social bonds that online communities use to commune around information disorders and how these social bonds are legitimated. The particular linguistic patterning of social bonds results in the formation of textual personae and can be illustrated via bond cluster diagrams. In terms of legitimation strategies, technological authority was the most common strategy discovered among the datasets, highlighting how the use of social media is important in the construal of credibility. Reflecting on the overall significance of the two case studies selected, this method was applicable to both non-political and political discourse, revealing similar discursive strategies as reflected by the creation of macro-categories. After explaining these key findings in further detail, the implications of these findings will be discussed in terms of applicability to studies in SFL and social semiotics, communication studies, and interdisciplinary research. Limitations of this research project and possible future research directions will then be reflected upon.

Broadly, this thesis has explored the interpersonal dimension of information disorders by focusing on how social bonds are legitimated. This has involved assembling an analytical framework that draws on both affiliation (Zappavigna, 2018; Zappavigna and Martin, 2018) and legitimation (Van Leeuwen, 2007), and which can be applied across visual, verbal, and written modes. The findings from this thesis based on the analysis of the multimodal social bonds in the YouTube videos and textual personae in the YouTube comments indicate the importance of understanding the key social bonds at stake in discourse in order to investigate how individuals and communities mobilise around shared values via specific communicative strategies designed to legitimate these values. It is only by identifying and then targeting these shared values that strategies

can be developed to counteract problematic and deceptive language. This aligns with current research in the field that emphasises that combatting information disorders requires more than just providing individuals or groups with logical arguments (Van der Linden and Roozenbeek, 2020; Vraga et al., 2019), rather focus needs to be directed towards how online communities are attached to specific values and interpersonal bonds with each other that help to strengthen the impact of information disorders. This chapter will now unpack this overview in further detail.

8.2. Key Findings

At a macro-level, the results of this thesis demonstrate the linguistic construction of social bonds and how they are legitimated across visual, verbal, and written modes within information disorder discourses. On social media as with other forms of mass media, individuals and organisations adapt their communication to appeal to their audiences and the conventions of the media platform used. Affiliation and legitimation are strategies for achieving these goals. The linguistic identification of affiliation and legitimation strategies complements the valuable research already undertaken regarding the impact of societal threats mediatised online and how they can be counteracted (Kashima et al., 2021; Nohrstedt, 2013; Wardle and Singerman, 2021).

The key findings of this thesis will be explored in four sections that correspond to the four key research questions: how affiliation and legitimation work multimodally across YouTube videos, the emergence of technological authority as a new key legitimation strategy on social media, textual personae as a method for understanding the complex array of identities that engage with information disorders online, and the inferences that can be made about how information disorders work across non-political versus political discourse based on the two case studies explored in this thesis. These key findings will now be further discussed.

8.2.1. YouTube Videos: Multimodal Affiliation and Legitimation

In answering the first research question, applying linguistic and multimodal methods meant that a systematic analysis could be conducted on the language and visual content of YouTube videos and comments in order to understand the interpersonal dimension of information disorders. This is important because people bond around shared values rather than via logical points alone (Vraga et al., 2019; Van der Linden and Roozenbeek, 2020). This approach contributed towards SFL and MDA studies by expanding these analytical frameworks to consider information disorders on YouTube, as well as contributing to other fields of study such as natural language processing and computational social sciences, that are interested in how manual analyses can be automated. Whilst this thesis does not claim that this analysis can be entirely automated (due to the complexity of such a task that involves knowledge of social context), it does hope to contribute to partially automating these sorts of analyses, as in an ever-changing world with torrents of information, automated analyses are needed to quickly understand large datasets. This can then be accompanied by closer qualitative analyses. These qualitative analyses can provide further insight into the specific strategies that YouTuber's adopt, with this information contributing to the development of guidelines of how to be more critically aware of deceptive and problematic information.

In Chapters 5 and 6, a multimodal analysis, that is, an analysis of meanings made in visual and verbal modes, was conducted on YouTube videos about the Momo Challenge and the Notre Dame Fire. This analysis revealed how YouTube videos relied on the construction of multimodal social bonds accompanied with multimodal legitimation strategies in order to generate persuasion. For example, the communing affiliation analysis conducted on the video transcripts showed which affiliation strategies were more frequently used in order to position audiences to align with particular values e.g., to persuade audiences that the Momo Challenge is real or create suspicion regarding the Notre Dame Fire. These strategies included addressing a community, emphasising particular social bonds, and confining a social bond to a limited set of options. The legitimation analysis then showed how these social bonds were positioned as credible

via particular linguistic choices (e.g. referring to statistics or recounting a story) or visual choices (e.g. showing a tweet or newspaper clippings). The purpose of this combined affiliation and legitimation analysis was to map out the key discursive themes across the video datasets and highlight the communicative strategies behind these persuasive videos that managed to garner interest from communities with various ideologies (as reflected by the analysis undertaken on the comments of these videos in Chapter 7).

Additionally, unique findings were linked to each case study. In the Momo Challenge case study (explored in Chapter 5), different macro-genres were identified encompassing news reporting, commentary, educational, entertainment, and click-bait. Despite this range of macro-genres, three key legitimation strategies appeared in each video. Firstly, references to news stories and experts were used to assert the newsworthiness (Bednarek and Caple, 2017) of the Momo Challenge, adopting rationalisation and authorisation strategies to portray the Momo Challenge as a serious issue. Visually, videos appeared as multi-layered documents (Bateman, 2008; Bateman, 2013) with simultaneous screens and banners, again, as a legitimation strategy to enhance the believability of the Momo Challenge. Secondly, intertextuality and evidence were key concerns across all the videos. As the particular dataset explored encompassed the most popular videos on the Momo Challenge, there were many references to other popular Momo Challenge videos, hence trying to establish a sense of credibility. Evidence in the form of screenshots and YouTube clips by both visual and verbal reference were also frequent, establishing a sense of technological authority. Lastly, all the videos in the Momo Challenge dataset semiotically enacted moral panics (Cohen, 1972). Legitimation strategies of mythopoesis were used in order to strengthen the moral panic, as well as the delegitimation of parents in effectively monitoring their children. Visually, what strengthened the impact of the moral panic was the striking image of "Momo", the salient image in the videos. The image of Momo acted as a bonding icon (Tann, 2012; Zappavigna and Martin, 2017) that the YouTubers could rally around in order to spread moral panics about Momo supposedly threatening children. These three key legitimation strategies regarding news stories and experts, intertextuality and evidence, and moral panics and mythopoesis, explain how the Momo

Challenge was able to become a viral sensation, capitalising on discourses around internet safety and children using technology.

The Notre Dame Fire case study incorporated a broader range of legitimation and social bonds working in conjunction with each other. In terms of macro-genres, the videos were either vlogs or voice-overs with live steam footage sharing conspiratorial and/or white supremacist discourses, and there was one instance of a news report sharing conspiratorial and white supremacist discourses. The key findings within this case study included the legitimation strategies of technological authority, mythopoesis, rationalisation, legitimating Euro-centric Christian views, and delegitimating authorities and conformity (in other words, mainstream thinking). In comparison to the Momo Challenge case study, whilst there was a wide array of legitimation and social bonds evident in the verbal content, the visual content was less salient. There were instances of videos that just consisted of a vlog format (a YouTuber speaking to the camera) without a reliance on visual content such as screenshots. Nonetheless, there were also examples that heavily relied on the visual representation of screenshots, tweets, and live stream footage in order to assert their claims that the Notre Dame Fire was a suspicious act. Whilst the visual content was more varied, there were instances in all the verbal content of the videos that alluded to screenshots, newspaper articles and YouTube as a platform. Hence, technological authority nevertheless had an important role across all the videos in legitimating the conspiratorial nature of the fire and delegitimating either politicians, immigrants, or Islam. Similar to the findings about the Momo Challenge and Momo as a bonding icon, there was one salient bonding icon that can be attributed to the intense interest in the Notre Dame Fire, that is, the image of the Notre Dame Cathedral on Fire. The Notre Dame Cathedral represents Western ideals, and this symbolism was frequently referred to throughout the video dataset, in order to rally people around hating immigrants and Islam, that YouTubers would portray as the cause of the fire.

8.2.2. Technological Authority

By investigating how social bonds are legitimated in the transcripts and visual content of YouTube videos that spread information disorder this illuminates the core values held by YouTubers and the linguistic and multimodal tools they use to legitimate these core values. This is useful because social bonds propagate information disorders, and it is important to analyse the specific social bonds held by different YouTubers, rather than homogenising the reasons why people engage in deceptive or conspiratorial discourses. Additionally, it is important to then understand how these social bonds are legitimated, as this explains why some social bonds are more impactful than others. This legitimation can occur via linguistic means (e.g. particular phrasings and emphases) or via multimodal means (e.g. via screenshots and website links). Thus, these means should be investigated in further detail to understand the affiliation potential of the content.

The framework for analysing social bonds and how they are legitimated across transcripts and visual content, can be applied to further studies that analyse videos and other multimodal content. A framework that considers the connection between social bonding and legitimation, avoids the analysis of an entity like a screenshot at face-value, but rather consider its broader context, in terms of its affiliation potential. Thus, this framework offers potential for systematically analysing language and visual content, a research area of interest for computational social science fields, but also considers the broader context of language and visual content, a research area of interest to linguistics and communication studies.

Across the analysis of the YouTube videos (Chapters 5 and 6) and YouTube comments (Chapter 7) there was one particularly pertinent legitimation strategy that should be reflected upon in further detail. This legitimation strategy was TECHNOLOGICAL AUTHORITY – how reference to the use of social media or other forms of media is used to construct credibility. Technological authority is an extension to the authorisation part of the legitimation framework. In the YouTube videos across the Momo Challenge (Chapter 5) and Notre Dame Fire (Chapter 6) case studies, visual references to

screenshots of social media posts, online articles, evidence collages, and split screen videos were common. In the Momo Challenge case study, there was the additional use of emulated screenshots in the YouTube videos, reflecting an interesting engagement with the affiliation potential of the screenshot as a believable artefact. In the Notre Dame Fire case study, screenshots and screen recordings played a crucial role in the conspiracies about the fire, however, there was a mix of videos that relied on verbal rather than visual references to screenshots and news articles in order to make their claims.

In the comment sections of both the Momo Challenge and Notre Dame Fire case studies (as explored in Chapter 7) technological authority was also an important legitimation strategy to validate shared social bonds. For example, in the Momo Challenge case study (Chapter 6), Connoisseurs, Sceptics, and Critics, each used strategies of technological authority. These were used to either legitimate the claim that the Momo Challenge is false by reference to internet cultures sites, or as a delegitimation strategy to question whether someone actually had the evidence (e.g., YouTube clips) that Momo is real. In the Notre Dame Fire case study, technological authority was a legitimation strategy mainly used in negative contexts as a strategy to spread false information. White Supremacists and Inciters used decontextualised screenshots to spread false information. In contrast, personae that shared truthful or unproblematic information about the fire such as Sceptics, Educators, and Confrontationists, used technological authority as a delegitimation strategy, questioning the veracity of screenshots and asking for further confirmation (in the form of direct links or expert opinions) to ascertain whether a digital artefact such as a video was real. Thus, technological authority was commonly used as both a legitimation and delegitimation tool. However, as this analysis shows, (de)legitimation strategies need to be considered in conjunction with an affiliation analysis in order to understand the social bonds that are being (de)legitimated, as this connection between social bonds and (de)legitimation strategies reveals the ideologies at stake in the information disorder sphere.

Overall, the exploration of technological authority as a concept demonstrates how YouTube videos can spread conspiratorial and hateful content in sophisticated ways,

and that digital artefacts such as screenshots contribute to the believability and virality of misinformation and disinformation. The addition of TECHNOLOGICAL AUTHORITY to the legitimation framework highlights the importance of technology as evidence in social media discourse. However, as explored by the Momo Challenge and Notre Dame Fire case studies, technological authority can be used in discourses that spread both truthful and false information, so technological authority needs to also be analysed in conjunction with an affiliation analysis that considers the key social bonds attached to the digital artefact. This study into technological authority also connected to the broader research currently being conducted on screenshots that encourages researchers to look at the screenshot itself as a media object or a kind of document, rather than simply looking 'through' the screenshot without considering its mediatisation and ignoring the societal importance of screenshots in our contemporary lives (Frosh, 2018).

8.2.3. Textual Personae: Affiliation and Legitimation

Identifying textual personae in the comment sections of YouTube videos that spread information disorders provides an understanding of how identities are linguistically enacted and how these various identities respond to information disorders. By being able to group textual personae according to social bonds and legitimation strategies this means that tailored responses to the issue of information disorder can be developed, that do not attempt to homogenise the multiple reasons why people engage in such hateful behaviour, thus allowing an understanding of the variety of social bonds and legitimation strategies at stake. The focus on textual personae also means that this method can be combined with other methods such as ethnography and qualitative interviews to gain further insight into these communities. The development of textual personae means that a shared linguistic terminology can be developed that allows greater communication amongst a wide array of disciplines, hence increasing the potential for interdisciplinary collaboration.

In Chapter 7, a combined affiliation and legitimation analysis was conducted on the comments of the YouTube videos discussed in Chapters 5 and 6. The overarching purpose of this chapter was to map out the different textual personae in YouTube

comment sections by investigating the core social bonds, affiliation and legitimation strategies unique to each persona. The results of this research demonstrate the significance of affiliation strategies as a means for persuading others to align with the sets of social bonds proposed in YouTube discourse, and how legitimation strategies help to further reinforce the social bonds shared by personae.

At a macro-level, we can think of the textual personae categories as representing the key ideational targets that were evaluated, how these ideational targets with affiliated with, and the strategies that were used to (de)legitimate the social bonds proposed. Identifying personae is a way of uniting all these different linguistic patterns to create a coherent profile of identity and ideology online. Furthermore, the (de)legitimation framework was useful in explaining the significance of the bonds that were identified, in the sense that the values personae affiliated with were also the main targets of their (de)legitimations. The dialogic affiliation analysis conducted on the replies to initiating comments of the videos provided insight into how values were interactively negotiated in response to various personae. This analysis revealed how the particular social bonds propagated by YouTubers are not always directly affiliated with, and instead may be challenged or modified in the comment feed in order to suit the agendas of particular commenters, including rejecting the bonds proposed in the video entirely.

The personae that were identified were mapped using bond cluster diagrams in order to show how the key bonds in their discourse were interrelated. These bond cluster diagrams visually encapsulate the key discoveries from the qualitative analysis and should not be read as a quantitative social network analysis. In the Momo Challenge case study, all personae either shared a 'safe' or 'dangerous' Momo bond, and a 'real' or 'fake' Momo bond. There were more crossovers in values across Momo Challenge personae, for example, the Critic personae shared both a 'fake Momo' and 'dangerous Momo' bond, thus providing a point of utility in an otherwise polarised discourse. The personae in the Notre Dame Fire dataset were even more polarised. Personae that shared problematic information about the Notre Dame Fire did not share any bonds with personae that shared unproblematic information about the Notre Dame Fire. This suggested how the absence of shared bonds can result in animosity, as there is no

common social bond (either positively or negatively evaluated) to use as a point of discussion and attempt to negotiate and affiliate around. In the Notre Dame Fire case study, there was also a greater array of ideational targets, compared to the Momo Challenge case study. This can be attributed to the politicised nature of the Notre Dame Fire case study where there is a greater array of competing ideologies, in comparison to the non-political nature of the Momo Challenge.

Overall, by considering different linguistic patterns I was able to identify various textual personae in the dataset. Additionally, by making the (de)legitimation framework more delicate, I could more clearly delineate textual personae who share problematic information from personae who share unproblematic information in terms of how they construe what counts as evidence. Thus, the bond cluster diagrams were useful for showing the 'macro-personae' categories that exist in terms of grouping personae, as well as indicating how each persona is unique in the particular bonds that they most commonly table.

8.2.4. Non-Political versus Political Discourse

By analysing the similarities and differences among non-politically motivated information disorders versus political motivated information disorders this means that further understandings towards the complexity of information disorders can be developed. The two case studies selected allowed an exploration of the spectrum of information disorders ranging from internet hoaxes, conspiracies, and white supremacist discourse. The virality of both case studies also means that their persuasiveness can be studied in detail via linguistic and multimodal methods. By applying this method to non-politically motivated and politically motivated case studies this shows the versatility of the methods adopted and their applicability to a range of different case studies. The results of this comparison hold significance for multiple disciplines that study information disorders by making clear how different forms of information disorders can be discussed and the similarities that emerge in terms of common linguistic and multimodal strategies used to spread information disorders.

After a close discourse analysis of the YouTube videos and comments associated with the Momo Challenge and the Notre Dame Fire case studies, some key inferences were made regarding non-politically motivated and politically motivated discourses: the shared role of social bonding around moral panics, the role of technological authority, and the development of macro-categories to encapsulate the range of textual personae discovered across multiple case studies. These inferences should not be seen as definitive findings, keeping in mind that only two case studies were explored, but rather as findings that illustrate how to compare and contrast vastly different case studies. These inferences can be made about the YouTube videos analysed (Chapters 5 and 6) and comments associated with these videos (Chapter 7).

In terms of the YouTube videos analysed, the concepts of social bonding around moral panics and technological authority were key. The non-politically motivated Momo Challenge YouTube videos included social bonds with the intent of spreading moral panics via the construction of a 'dangerous Momo bond', as well as intertextual references (which was absent with the Notre Dame Fire videos) that linked to the social bonds shared in other Momo Challenge YouTube videos. The politically motivated Notre Dame Fire case study also featured social bonds that spread a moral panic, however, these social bonds featured a wider range of ideational targets, with examples including 'evil Muslims bond', 'evil Elites bond', 'dangerous Direct Energy Weapons bond', and 'evil Islam bond'. In both the Momo Challenge and Notre Dame Fire case studies, as explored in the previous section, technological authority was a key legitimation and delegitimation strategy. In the videos across both case studies, split screen video formats, and a reliance of screenshots and live streams were prominent visual strategies.

The analysis of YouTube comments revealed additional insights by comparing the textual personae in both political and non-political discourse. Firstly, six macro-categories were identified by establishing which personae shared similar evaluation strategies and communicative goals. For example, Critics and Confrontationists both negatively evaluated individuals in order to assert the authenticity of a YouTube video or individual. Secondly, there were also differences in which personae were represented.

For example, in the Momo Challenge case study Nationalists were identified, but were absent in the Notre Dame Fire case study. This was attributed to the Notre Dame Fire commenters not making their nationality evident and instead targeting non-Western cultures and individuals, and aligning instead with particular religions and races. Thirdly, the Notre Dame Fire case study featured more distinct types of personae sharing false information (Anti-Elitists, Anti-Media, White Supremacists, Inciters) compared to the Momo Challenge case study (Myth Spreaders and Moralisers). This was due to the Notre Dame Fire having a broader range of negatively evaluated ideational targets (i.e., the targets of hate speech and vast political ideologies engaged with) compared to the Momo Challenge, that perhaps due to its unpolitical nature, had fewer ideational targets as the debate in the comment threads revolved around hearsay and moral evaluations of the internet, rather than targeting particular individuals or political groups. These were the key points of similarities and differences discovered across the Momo Challenge and Notre Dame Fire comments analysis.

Overall, based on the observations discussed in this section, there were more similarities than differences among the case studies, as reflected by the macro-personae categories identified and the concept of technological authority. However, as discussed, if delving deeper in specific personae, or considering further sub-types of technological authority as a legitimation and delegitimation strategy, there were distinctions among the case studies. This research has reflected on political versus non-political discourse in the context of information disorders, however, further case studies would need to be explored to develop more specific qualitative or quantitative findings.

8.3. Implications

This section will discuss in further detail the implications of this thesis across several different areas. Firstly, the contributions of this research towards Systemic Functional Linguistics (SFL) and Social Semiotics will be discussed, in terms of how the frameworks used in this thesis expand upon our understandings of the linguistic and multimodal dimensions of misinformation and disinformation. Secondly, the relevance of this research to communication studies (broadly conceived as research interested in human

communication) will be discussed, particularly with regard to how the research in this thesis can contribute to understanding communities involved in misinformation and disinformation in further detail. Lastly, this section will explore how this thesis engages with interdisciplinary research and has further potential to engage across disciplines such as computational social sciences, social psychology, and natural language processing.

8.3.1. Systemic Functional Linguistics and Social Semiotics

This thesis has implications for how affiliation and legitimation are modelled in SFL and multimodally in social semiotics. The affiliation framework has been augmented in this study to consider social bonds at a more macro-level. This was achieved by generalising the key ideational and attitudinal targets, and affiliation strategies expressed in the discourse as personae categories that were illustrated in the bond cluster diagrams in Chapter 7. In addition, the communing affiliation framework has been analysed and applied in detail in order to account for how social bonds are broadened to consider multiple options or confined to one option (finessing), how vocatives are used to rally others or how names are used to rally specific communities (convoking), and how specific typographical, numerical choices, or superlatives are used to emphasise claims (tempering). Traditional SFL theorists might question why communing affiliation has been the focus of this thesis rather than studies of grammar (like transitivity, mood and theme) e.g., the phrase *'here in the Philippines'* could be annotated as a marked theme, however, this thesis annotates it as having a designating affiliative function. The reason for this choice is that this thesis focuses on ascribing interpersonal meaning at a discourse semantic level in digital environments, in other words, understanding what particular linguistic strategies propagate the social bonds identified. The affiliation framework is being continuously adapted, with a new affiliation framework detailed in Doran et al. (forthcoming) that broadens affiliation from the discourse semantic level to the tenor level. This thesis provides a small contribution to thinking about the affiliation framework in regards to macro social bonds, the mapping of textual personae, and legitimation strategies, which will aid future SFL researchers in applying, visualising, and adapting the affiliation framework further.

Another key contribution of this thesis to studies in SFL and social semiotics, is the extension of the legitimation framework. This thesis refined the 'authority' sub-category to a more delicate level in order to capture the specific ways in which impersonal authority is construed. Two new additions were made to authority: 'marketing', credibility shown via emblems such as company logos, and 'technological', credibility shown via technological means e.g., screenshots, online articles, references to video links or google searches. This has applications across written, verbal and visual content analysis in multimodal and social semiotic studies.

8.3.2. Communication Studies

The methodology and results from this thesis are also applicable to research into human communication regarding information disorders and can be applied to a range of different case studies and applications. By being able to linguistically identify the different social values and affiliations strategies used in discourse, tailored responses to the issues of information disorder and hate speech can be developed, that do not attempt to homogenise the multiple reasons why people engage in such hateful behaviour (Inwood and Zappavigna, 2022a). The homogenisation of information disorders is dangerous because it fails to capture the unique socio-cultural circumstances of online communities and how they interact with information disorders. The focus on textual personae enables a linguistic profile to be developed for specific communities, intersecting with other methods in communication studies such as ethnography and qualitative interviews.

Additionally, the results of this study relate to previous studies on the mediatization of societal threats (Krzyżanowski et al., 2018) by highlighting the power of rhetoric in 'othering' certain groups in order to gain popularity and emphasising how 'the politics of fear' (Wodak, 2015) works at a discourse semantic level. Therefore, the results of this study engage with this previous valuable research but include the added dimension of affiliation in order to explore how social values that encourage fear of the 'other' are bonded around and legitimated. Again, this research is useful to research in communication studies because it can provide a linguistic and multimodal dimension

to concepts that are a key concern in communication studies such as societal threats and information disorders.

The focus of SFL research into studying white supremacy and extremist communication can also be useful in designing educational materials so that students can linguistically identify the language used by white supremacists and extremists and prevent themselves from becoming radicalised, as discussed by Szenes (2021). Understanding the importance of evaluative language and stance in white supremacist and conspiratorial discourse is an increased area of focus by researchers working across a range of discourse analysis perspectives (Szenes, 2021; Demata et al., 2022). This increased interest shows a need to further delineate hateful communities according to these linguistic strategies and collaboratively design materials that educate a diverse range of populations. To summarise, methods in SFL intersect with key issues raised in communication and media literacy studies, that highlight the complexity in effectively communicating the multi-layered issues at hand and translating academic results to varied audiences.

8.3.3. Interdisciplinary Research

This thesis is also relevant to broader qualitative and quantitative studies that focus on societal threats and information disorders. Its method, rather than a single solution to solving information disorders, should be seen as an avenue with potential to connect research from multiple disciplines together. For example, understanding the personae in comment sections can aid in the creation of manual and automated detection processes for identifying the specific networks of values that charge white supremacy and conspiratorial discourse. This approach might expand upon research already conducted on YouTube comments, the majority of which is quantitative (Allington and Joshi, 2020; Miller, 2021; Röchert et al., 2021). These studies, however, do not engage with the concept of textual personae, despite its potential to afford more accessible understandings of the construction of information disorders and to aid strategies for counteracting this dangerous language. The notion of personae allows us to start unpacking the distinct linguistic profiles of a variety of communities with an array of

values, and also to integrate different research methods through making explicit shared linguistic terminology.

With the recent popularity of language generation models like GPT-3 (Generative Pretrained Transformer 3) and ChatGPT developed by OpenAi (the version of GPT-3 fine-tuned for conversational language) it is now possible for anyone to give ChatGPT a series of YouTube comments and ask it to categorise these comments according to linguistically enacted identities. Nonetheless, ChatGPT's current limitations are in its lack of contextual information and inability to decipher truthful versus false information as it relies on the patterns it has learned from the data it was trained on and the specific prompts given. ChatGPT has not been trained in applied linguistic methods like SFL and so even if it does generate a list of linguistically enacted identities it is based on common-sense terms rather than a systematic approach to understanding the particular language features of a YouTube comment. Additionally, ChatGPT would not be able to generate the bond cluster diagrams developed in this thesis without further contextual information and knowledge of social bonding. However, ChatGPT in its current form does offer the ability to help with sampling strategies, for example, quickly categorising YouTube comments into shared themes and language structures (e.g. comments with hateful language). This would then provide applied linguists with the ability to conduct a manual analysis on comments representative of a larger dataset.

This thesis also provides a way of delving deeper into the social values of conspiratorial communities identified by computational social science studies (such as (Shahsavari et al., 2020) by highlighting the affiliation strategies used to make these social values prominent. Therefore, it may be of value to quantitative researchers aiming to understand the affiliative strategies of the most commonly identified social values in their datasets, or to qualitative researchers wanting to add another dimension of coding to their data by using linguistic terminology to articulate how communities are addressed and social values are rejected or rallied around. Whilst this thesis has only focused upon two case studies, interdisciplinary approaches might expand into other domains such as health misinformation.

8.4. Limitations and Future Directions

The methods adopted in this thesis do have several drawbacks. Firstly, the intensive manual analysis conducted on the transcripts, visual content, and comments is difficult to scale to larger datasets. Simplifying or partially-automating the analysis could result in a loss of reliability and analytic detail. For this reason, the results remain qualitative in this thesis and thus may not meet positivist understandings of generalizability. In addition, the approach adopted in this thesis should not be viewed as a single solution to combatting issues of conspiracy and white supremacy, rather it is intended to work in conjunction with other approaches in order to leverage the power of interdisciplinary insights.

However, there are many possibilities for how this research could be expanded upon in the future which may remedy some of these limitations. Firstly, it would be interesting to consider which personae categories are portable enough to apply to multiple case studies, and which categories are dependent on the particular case study context. Future research, in applying the methods in this thesis to domains such as misinformation about health information, might consider why some kinds of personae span across multiple contexts. This work might also refine the analytical approach to more closely identify the patterns in language relevant to these potential domains, and further specify relationships between affiliation strategies, legitimation strategies, and bonds. Additionally, the visual analysis of this research could be complemented with further studies that interview users who engage with screenshots and question their motives for creating and sharing screenshots, particularly in breaking news contexts. Whilst this thesis has remained explorative since it has focused on the application of new methods, rather than providing specific frequency information regarding the affiliation and legitimation strategies explored across large datasets, future studies might attempt to quantify these patterns across visual and verbal modes and consider the extent to which the visual and the verbal meanings coordinate with each other.

In terms of research impact, more collaboration could occur with computational linguists, and theorists in computational social science and media disciplines to make

the findings of this study more accessible for a broader audience. Tailored responses to issues of information disorder could involve creating a linguistic profile for a specific community, outlining the social values and affiliative strategies particular to that community, and then developing strategies that can counteract this hateful discourse. These strategies need the input of both researchers and people who have lived experience in the communities studied in order to develop solutions. As Mahl et al. (2022: 18) highlight, "national, cultural, and socio-political contexts" need to be taken into account when studying conspiracy theories and extremist discourse. This means that "comparative studies across topics" also need to be developed if broader claims about conspiratorial communities and information disorders are to be made. Thus, this research is one small contribution to the great interdisciplinary effort that is needed in order to counteract the significant social harm that is currently being enacted through the proliferation of information disorders and hate speech online.

9. REFERENCE LIST

Abraham B (2014) Challenging Hate Speech With Facebook Flarf: The Role of User Practices in Regulating Hate Speech on Facebook. *Fibreculture Journal: Internet Theory Criticism Research*. (23): 47-72.

Agha A (2005) *Registers of Language.* Malden: John Wiley & Sons.

Allcott H and Gentzkow M (2017) Social media and fake news in the 2016 election. *Journal of Economic Perspectives* 31(2): 211-236.

Allington D, Buarque BL and Barker Flores D (2021) Antisemitic conspiracy fantasy in the age of digital media: Three 'conspiracy theorists' and their YouTube audiences. *Language and Literature* 30(1): 78-102.

Allington D and Joshi T (2020) "What Others Dare Not Say": An Antisemitic Conspiracy Fantasy and its YouTube Audience. *Journal of Contemporary Antisemitism* 3(1).

Amarasingam A (2014) *The Stewart/Colbert effect: Essays on the real impacts of fake news.* Jefferson: McFarland.

Andre V (2012) 'Neojihadism'and YouTube: Patani militant propaganda dissemination and radicalization. *Asian Security* 8(1): 27-53.

Anthony LA (2014) Computer software. *Tokyo, Japan: Waseda University*.

Asakitikpi AO and Gadzikwa J (2020) A Critical Discourse Analysis of Online Youtube News Coverage of South African Discourses on Xenophobia in Democratic South Africa. *Politikon* 47(4): 479-493.

Bae SS and Baxter S (2018) YouTube videos in the English language as a patient education resource for cataract surgery. *International Ophthalmology* 38(5): 1941-1945.

Bakir V and McStay A (2018) Fake news and the economy of emotions: Problems, causes, solutions. *Digital Journalism* 6(2): 154-175.

Barnes JA (1994) *A Pack of Lies: Towards a Sociology of Lying.* Cambridge: Cambridge University Press.

Bartlett J and Miller C (2010) *The Power of Unreason: Conspiracy Theories, Extremism and Counter-Terrorism.* London: Demos.

Basch CH, Mongiovi J, Berdnik A and Basch CE (2016) The most widely viewed YouTube videos with content related to multivitamins. *Health Promotion Perspectives* 6(4): 213-216.

Bateman J (2008) *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents.* Hampshire: Springer.

Bateman J, McDonald D, Hiippala T, Couto-Vale, D and Costezki, E. (2019) Systemic-functional linguistics and computation: new directions, new challenges. In G. Thompson, WL Bowcher, L Fontaine, JY Liang (eds) *The Cambridge Handbook of Systemic-Functional Linguistics*. Cambridge: Cambridge University Press.

Bateman J, Wildfeuer J and Hiippala T (2017) *Multimodality: Foundations, research and analysis–A problem-oriented introduction.* Berlin: Walter de Gruyter.

Bateman JA (2013) Multimodal Analysis of Film within the GeM framework. *Ilha do Desterro A Journal of English Language, Literatures in English and Cultural Studies* (64): 49-84.

Bavelas JB, Black A, Chovil N and Mullett J (1990) Truths, lies, and equivocations: The effects of conflicting goals on discourse. *Journal of Language and Social Psychology* 9(1-2): 135-161.
Bednarek M (2006) *Evaluation in Media Discourse: Analysis of a Newspaper Corpus.* London: A&C Black.
Bednarek M (2014) The television title sequence: a visual analysis of *Flight of the Conchords*. In E Djonov and S Zhao (eds) *Critical Multimodal Studies of Popular Culture*. London/New York: Routledge: 36-54.
Bednarek M and Caple H (2014) Why do news values matter? Towards a new methodological framework for analysing news discourse in Critical Discourse Analysis and beyond. *Discourse & Society* 25(2): 135-158.
Bednarek M and Caple H (2017) *The discourse of news values: How news organizations create newsworthiness.* Oxford: Oxford University Press.
Bennett WL and Livingston S (2018) The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication* 33(2): 122-139.
Benson P (2016) *The Discourse of YouTube : Multimodal Text in a Global Context.* New York: Routledge.
Bernstein B (2003) *Class, codes and control: Applied studies towards a sociology of language.* London: Psychology Press.
Biber D and Finegan E (1989) Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text-interdisciplinary journal for the study of discourse* 9(1): 93-124.
Biddle WW (1931) A psychological definition of propaganda. *The Journal of Abnormal and Social Psychology* 26(3): 283-295.
Bigman CA, Planey AM, Williamson LD, et al. (2022) 'There will be screen caps': the role of digital documentation and platform collapse in propagation and visibility of racial discourses. *Information, Communication & Society*. 1-18.
Bishop S (2018) Anxiety, panic and self-optimization: Inequalities and the YouTube algorithm. *Convergence* 24(1): 69-84.
Bok SJP (1978) *Lying: Moral choice in private and public life.* Britain: The Harvester Press Limited.
Boler M and Turpin E (2008) The Daily Show and Crossfire: Satire and sincerity as truth to power. In: Boler M (ed) *Digital Media and Democracy: Tactics in Hard Times*. Cambridge, United States: MIT Press, pp.383-404.
Bolíbar M (2016) Macro, meso, micro: broadening the 'social' of social network analysis with a mixed methods approach. *Quality & Quantity* 50(5): 2217-2236.
Bradac JJ (1983) The language of lovers, flovers, and friends: Communicating in social and personal relationships. *Journal of Language and Social Psychology* 2(2-3-4): 141-162.
Brewer PR, Young DG and Morreale M (2013) The impact of real news about "fake news": Intertextual processes and political satire. *International Journal of Public Opinion Research* 25(3): 323-343.
Brummette J, DiStaso M, Vafeiadis M, et al. (2018) Read All About It: The Politicization of "Fake News" on Twitter. *Journalism & Mass Communication Quarterly* 95(2): 497-517.

Bruns A and Moe H (2014) Structural layers of communication on Twitter. In A Bruns , M Mahrt, K Weller, J Burgess and C Puschmann (Eds) *Twitter and Society [Digital Formations, Volume 89]*. New York: Peter Lang, pp.15-28.

Burgess J and Green J (2018) *YouTube: Online video and participatory culture.* New Jersey: John Wiley & Sons.

Burgoon JK, Buller DB, Guerrero LK, et al. (1996) Interpersonal deception: XII. Information management dimensions underlying deceptive and truthful messages. *Communication Monographs* 63(1).

Busby J (2017) Social media and the scholar in an era of hyper-Nationalism and fake news. *PS: Political Science & Politics* 50(4): 1004-1007.

Caldwell D (2010) Making many meanings in popular rap music. In A Mahboob and NK Knight (Eds) *Appliable Linguistics*. London: Bloomsbury, pp. 234-250.

Caple H (2008) Intermodal relations in image nuclear news stories. In L Unsworth (ed) *Multimodal semiotics: Functional analysis in contexts of education*. London: Bloomsbury, 125-138.

Carlson M (2018) Fake news as an informational moral panic: the symbolic deviancy of social media during the 2016 US presidential election. *Information, Communication & Society*. 1-15.

Carlson M (2020) Fake news as an informational moral panic: the symbolic deviancy of social media during the 2016 US presidential election. *Information, Communication & Society* 23(3): 374-388.

Cassam Q (2019) *Conspiracy theories.* Cambridge: Polity.

Castelfranchi C and Poggi I (1994) Lying as pretending to give information. In H Parret (Ed) *Pretending to Communicate*. Berlin: De Gruyter: pp. 276-291.

Chaidas D (2018) The benefits of narratology in the analysis of multimodal legitimation: The case of New Democracy. *Discourse & Communication* 12(3): 258-277.

Chiluwa I and Ifukor P (2015) 'War against our Children': Stance and evaluation in# BringBackOurGirls campaign discourse on Twitter and Facebook. *Discourse & Society* 26(3): 267-296.

Chisholm RM and Feehan TD (1977) The intent to deceive. *The Journal of Philosophy* 74(3): 143-159.

Christie F (1999) Genre theory and ESL teaching: A systemic functional perspective. *Tesol Quarterly* 33(4): 759-763.

Christie F and Derewianka B (2010) *School discourse: Learning to write across the years of schooling.* London: A&C Black.

Christie F and Martin JR (2009) *Language, knowledge and pedagogy: Functional linguistic and sociological perspectives.* London: Bloomsbury Publishing.

Clarke I and Grieve J (2017) Dimensions of abusive language on Twitter. *Proceedings of the First Workshop on Abusive Language Online.* 1-10.

Cléirigh C (1998) *A selectionist model of the genesis of phonic texture: systemic phonology and universal Darwinism*. Unpublished PhD Thesis, University of Sydney.

Coady D (2012) *What to believe now: Applying epistemology to contemporary issues.* West Sussex: John Wiley & Sons.

Cohen S (1972) *Folk Devils and Moral Panics: The Creation of the Mods and Rockers.* Oxford: Martin Robertson.

Coltman-Patel T, Dance W, Demjén Z, Gatherer D, Hardaker C and Semino E (2022) 'Am I being unreasonable to vaccinate my kids against my ex's wishes?'–A corpus linguistic exploration of conflict in vaccination discussions on Mumsnet Talk's AIBU forum. *Discourse, Context & Media* 48: 1-9.

Corry F (2021) Screenshot, save, share, shame: Making sense of new media through screenshots and public shame. *First Monday* 26(4).

Coupland N (2002) *Language, situation and the relational self: Theorizing dialect-style in sociolinguistics.* Cambridge: Cambridge University Press.

Culloty E (2020) Conspiracy and the epistemological challenges of mediatized conflict. In S Maltby, B O'Loughlin, K Parry and L Roselle (eds) *Spaces of War, War of Spaces*. London: Bloomsbury, 83-102.

d'Ancona M (2017) *Post-truth: The new war on truth and how to fight back.* London: Ebury Press.

D'Onofrio A (2018) Personae and phonetic detail in sociolinguistic signs. *Language in Society* 47(4): 513-539.

Dahl T and Fløttum K (2017) Verbal–visual harmony or dissonance? A news values analysis of multimodal news texts on climate change. *Discourse, Context & Media* 20: 124-131.

Daneš F (1964) A three-level approach to syntax. *Travaux linguistiques de Prague* (1), 225-240.

Daniels J (2009) *Cyber racism: White supremacy online and the new attack on civil rights.* Washington DC: Rowman & Littlefield Publishers.

Dare-Edwards HL (2014) 'Shipping bullshit': Twitter rumours, fan/celebrity interaction and questions of authenticity. *Celebrity Studies* 5(4): 521-524.

De Saussure F (1916) Nature of the linguistic sign. *Cours de linguistique generale*. Paris: Payot.

de Unamuno M (1924) *Essays and soliloquies.* New York: AA Knopf.

Demata M, Zorzi V and Zottola A (2022) Conspiracy theory discourse: critical inquiries into the language of anti-science, post-trutherism, mis/disinformation and alternative media. In: Demata M, Zorzi V and Zottola A (eds) *Conspiracy Theory Discourses*. Amsterdam: John Benjamins.

Dictionaries O (2001) *Paperback Oxford English Dictionary.* Oxford: Oxford University Press.

Dictionary OE (nd) *misinformation | Origin and meaning of misinformation by Online Etymology Dictionary*. Available at: https://www.etymonline.com/word/misinformation#etymonline_v_44758. Accessed 1st Feb 2023.

Dictionary OE (nd) *Oxford English Dictionary*. Oxford: Oxford University Press, Available at: https://en.oxforddictionaries.com/. Accessed 1st Feb 2023.

DiFonzo N and Bordia P (2007) Rumor, gossip and urban legends. *Diogenes* 54(1): 19-35.

Djonov E (2005) *Analysing the organisation of information in websites: From hypermedia design to systemic functional hypermedia discourse analysis.* Unpublished PhD Thesis, University of New South Wales.

Don A (2007) *A Framework for the investigation of interactive norms and the construction of textual identity in written discourse communities: The case of an email discussion list*. Unpublished PhD Thesis, University of Birmingham.

Don A (2009) Legitimating tenor relationships: Affiliation and alignment in written interaction. *Linguistics & the Human Sciences* 5(3).
Don A (2016) "It is hard to mesh all this": Invoking attitude, persona and argument organisation. *Functional Linguistics* 3(1): 1-26.
Don A (2018) Stance-taking and the construal of textual persona in written contexts: Social Contact revisited. *Linguistics and the Human Sciences* 13(1-2): 70-95.
Donzelli G, Palomba G, Federigi I, et al. (2018) Misinformation on vaccination: A quantitative analysis of YouTube videos. *Human Vaccines & Immunotherapeutics* 14(7): 1654-1659.
Doran Y (2017) *The Discourse of Physics: Building Knowledge through Language, Mathematics and Image.* London: Routledge.
Doran Y, Martin JR and Zappavigna M (forthcoming) *Affiliation and Systemic Functional Linguistics: Negotiating Community.* Sheffield: Equinox.
Dow BJ, Johnson AL, Wang CS, Whitson J and Menon T. (2021) The COVID-19 pandemic and the search for structure: Social media and conspiracy theories. *Social and Personality Psychology Compass* 15(9): 1-22.
Drasovean A and Tagg C (2015) Evaluative Language and Its Solidarity-Building Role on TED.com: An Appraisal and Corpus Analysis. *Language@Internet* 12.
Dreyfus SJ, Humphrey S, Mahboob A and Martin JR. (2015) *Genre pedagogy in higher education: The SLATE project.* London: Palgrave Macmillan.
du Preez A and Lombard E (2014) The role of memes in the construction of Facebook personae. *Communicatio* 40(3): 253-270.
Dynel M (2014) Participation framework underlying YouTube interaction. *Journal of Pragmatics* 73: 37-52.
Eggins S (2004) *Introduction to Systemic Functional Linguistics.* London: A&C Black.
Eggins S and Slade D (2005) *Analysing Casual Conversation.* Sheffield: Equinox Publishing Ltd.
Ekman M (2014) The dark side of online activism: Swedish right-wing extremist video activism on YouTube. *MedieKultur: Journal of media and communication research* 30(56): 79-98.
Ekman P (1985) *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage.* New York: W.W. Norton & Company.
Erkama N and Vaara E (2010) Struggles over legitimacy in global organizational restructuring: A rhetorical perspective on legitimation strategies and dynamics in a shutdown case. *Organization Studies* 31(7): 813-839.
Ernst J, Schmitt JB, Rieger D, et al. (2017) Hate beneath the counter speech? A qualitative content analysis of user comments on YouTube related to counter speech videos. *Journal for Deradicalization*.(10): 1-49.
Fallis D (2014) The Varieties of Disinformation. In L Floridi and P Illari (Eds) *The Philosophy of Information Quality*. Synthese Library, vol 358. Cham: Springer, pp.135-161.
Fallis D (2015) What is disinformation? *Library Trends* 63(3): 401-426.
Farkas J and Schou J (2018) Fake News as a Floating Signifier: Hegemony, Antagonism and the Politics of Falsehood. *Javnost - The Public* 25(3): 298-314.
Farkas J, Schou J and Neumayer C (2018) Platformed antagonism: racist discourses on fake Muslim Facebook pages. *Critical Discourse Studies* 15(5): 463-480.

Feez S, Iedema R and White P (2008) *Media literacy*. Sydney: NSW Adult Migrant Education Service.
Feng D and O'Halloran KL (2013) The multimodal representation of emotion in film: Integrating cognitive and semiotic approaches. *Semiotica* 2013(197): 79-100.
Finlayson A (2020) YouTube and Political Ideologies: Technology, Populism and Rhetorical Form. *Political Studies 70(1): 62-80*.
Finneman T and Thomas RJ (2018) A family of falsehoods: Deception, media hoaxes and fake news. *Newspaper Research Journal* 39(3): 350-361.
Firbas J (1964) On defining the theme in functional sentence analysis. *Linguistiques de Prague* 1: 267-280.
Firth JR (1950) Personality and language in society. *The Sociological Review* 42(1): 37-52.
Firth JR (1957) *Papers in linguistics, 1934-1951.* Oxford: Oxford University Press.
Floridi L (1996) Brave.Net.World: the Internet as a disinformation superhighway? *The Electronic Library* 14(5): 8.
Flowerdew J and Wan A (2010) The linguistic and the contextual in applied genre analysis: The case of the company audit report. *English for Specific purposes* 29(2): 78-93.
Fong A, Roozenbeek J, Goldwert D, et al. (2021) The language of conspiracy: A psychological analysis of speech used by conspiracy theorists and their followers on Twitter. *Group Processes & Intergroup Relations* 24(4): 606-623.
Foster EK (2004) Research on gossip: Taxonomy, methods, and future directions. *Review of general psychology* 8(2): 78-99.
Foucault M (2021) *Language, counter-memory, practice.* New York: Cornell University Press.
Fox CJ (1983) *Information and Misinformation. An Investigation of the Notions of Information, Misinformation, Informing, and Misinforming.* Westport, CT and London.: Greenwood Press.
Frank MG and Svetieva E (2013) Deception. In: D Matsumoto, MG Frank and HS Hwang (eds) *Nonverbal communication: Science and applications*. Thousand Oaks, CA: Sage, pp. 114-132.
Franzke AS, Bechmann A, Zimmer M and Ess C (2020) Internet Research: Ethical Guidelines 3.0. *Association of Internet Researchers*. https://aoir.org/reports/ethics3.pdf Accessed 1st Feb 2023.
Frosh P (2018) *The Poetics of Digital Media.* New Jersey: John Wiley & Sons.
Fuchs C (2018) "Dear Mr. Neo-Nazi, Can You Please Give Me Your Informed Consent So That I Can Quote Your Fascist Tweet?" Questions of Social Media Research Ethics in Online Ideology Critique In G Meikle (Ed) *The Routledge Companion to Media and Activism*. London: Routledge, pp. 385-394.
Fuchs C (2019) *Nationalism on the Internet: Critical theory and ideology in the age of social media and fake news.* London: Routledge.
Galasinski D (2018) Lying and Discourse Analysis. In J Meibauer (ed) *The Oxford Handbook of Lying*. Oxford: Oxford University Press, pp.516-528.
Giasemidis G, Singleton C, Agrafiotis I, Nurse JRC, Pilgrim A, Willis C and Greetham DV (2016) Determining the Veracity of Rumours on Twitter. In E Spiro and YY Ahn (Eds) *Social Informatics*. SocInfo 2016. Lecture Notes in Computer Science, vol 10046 Cham: Springer International Publishing, 185-205.

Giglietto F, Rossi L and Bennato D (2012) The open laboratory: Limits and possibilities of using Facebook, Twitter, and YouTube as a research data source. *Journal of Technology in Human Services* 30(3-4): 145-159.

Gillborn D (2006) Rethinking white supremacy: Who counts in 'WhiteWorld'. *Ethnicities* 6(3): 318-340.

Glaser Barney G and Strauss Anselm L (1967) *The discovery of grounded theory: strategies for qualitative research*. New York: Adline de Gruyter.

Glenski M, Ayton E, Mendoza J and Volkova S (2019) Multilingual Multimodal Digital Deception Detection and Disinformation Spread across Social Platforms. *arXiv preprint arXiv:1909.05838*.

Goffman E (1978) *The Presentation of Self in Everyday Life.* Edinburgh: Edinburgh University Press.

Golovchenko Y, Buntain C, Eady G, Brown MA and Tucker JA. (2020) Cross-Platform State Propaganda: Russian Trolls on Twitter and YouTube During the 2016 US Presidential Election. *The International Journal of Press/Politics* 25(3) 357-389.

Green B (2021) US Digital Nationalism: A Habermasian Critical Discourse Analysis of Trump's 'Fake News' Approach to the First Amendment. In A MacKenzie, J Rose, I Bhatt (Eds) *The Epistemology of Deceit in a Postdigital Era: Postdigial Science and Education*. Cham: Springer, 95-117.

Grice HP (1975) Logic and conversation. In P Cole and JL Morgan (eds) *Syntax and Semantics: Speech Acts*. New York: Academic Press, pp.41-58.

Gupta A, Kumaraguru P, Castillo C and Meier P (2014) TweetCred: Real-Time Credibility Assessment of Content on Twitter. In LM Aiello and D McFarland (eds) *Social Informatics: 6th International Conference, SocInfo 2014, Barcelona, Spain, November 11-13, 2014. Proceedings*. Cham: Springer International Publishing, pp.228-243.

Gupta S, Sakamoto K and Ortony A (2012) Telling it like it isn't: a comprehensive approach to analyzing verbal deception. In F Paglieri, L Tummonlini, R Falcone and M Micceli (Eds) *The goals of cognition: Festschrift for Cristiano Castelfranchi*. London: College Publications 579-612.

Hadley G (2017) *Grounded theory in applied linguistics research: A practical guide.* London: Taylor & Francis.

Haiden L and Althuis J (2018) The Definitional Challenges of Fake News. King's Centre for Strategic Communications, Department of War Studies, King's College London. *International Conference on Social Computing, Behaviour-Cultural Modelling, and Prediction and Behaviour Representation in Modelling and Simulation*, Washington.

Halliday MAK (1961) Categories of the theory of grammar. *Word* 17(3): 241-292.

Halliday MAK (1978) *Language as social semiotic: The social interpretation of language and meaning.* London: Hodder Arnold.

Halliday MAK (1995) *Discourse in society: Systemic functional perspectives.* Westport: Greenwood Publishing Group.

Halliday MAK (2008) Working with meaning: Towards an appliable linguistics. In JJ Webster (Ed) *Meaning in context: Implementing intelligent applications of language studies*. London: Continuum. 7-23.

Halliday MAK and Hasan R (1989) Language, context, and text: Aspects of language in a social-semiotic perspective. Oxford: Oxford University Press.

Halliday MAK and Matthiessen C (1999) *Construing experience through meaning: A language-based approach to cognition.* London: Bloomsbury Publishing.
Halliday MAK and Matthiessen CM (2013) *Halliday's Introduction to Functional Grammar.* London: Routledge.
Hanson G and Haridakis P (2008) YouTube users watching and sharing the news: A uses and gratifications approach. *Journal of Electronic Publishing* 11(3).
Hao J and Humphrey S (2009) The Role of'Coupling'in Biological Experimental Reports. *Linguistics & the Human Sciences* 5(2).
Hardaker C (2010) Trolling in asynchronous computer-mediated communication: From user discussions to academic definitions. *Journal of Politeness Research* 6(2) 215-242.
Harper D (n.d.) Etymology Dictionary. *Online Etymology Dictionary.* Accessed 1st Feb 2023.
Hart C (2017) Metaphor and intertextuality in media framings of the (1984–1985) British miners' strike: A multimodal analysis. *Discourse & Communication* 11(1): 3-30.
Hasan R (1989) Semantic variation and sociolinguistics. *Australian Journal of Linguistics* 9(2): 221-275.
Heley K, Gaysynsky A and King AJ (2022) Missing the Bigger Picture: The Need for More Research on Visual Health Misinformation. *Science communication* 44(4): 514-527.
Hemsley J and Snyder J (2018) Dimensions of visual misinformation in the emerging media landscape. *Misinformation and mass audiences* 91: 2018.
Hermida A (2010) Twittering the news: The emergence of ambient journalism. *Journalism practice* 4(3): 297-308.
Hjelmslev L (2019) *La stratification du langage.* Berlin: De Gruyter Mouton.
Hokka J (2021) PewDiePie, racism and Youtube's neoliberalist interpretation of freedom of speech. *Convergence* 27(1): 142-160.
Holan A (2017) The media's definition of fake news vs. Donald Trump's. *First Amendment Law Review 16:* 121-128.
Holdcroft D (1991) *Saussure: signs, system and arbitrariness.* Cambridge: Cambridge University Press.
Hood S (2010) *Appraising Research: Evaluation in Academic Writing.* London, UK: Palgrave Macmillan.
Hood S and Lander J (2016) Technologies, modes and pedagogic potential in live versus online lectures. *International Journal of Language Studies* 10(3).
Hopper R and Bell RA (1984) Broadening the deception construct. *Quarterly Journal of Speech* 70(3): 288-302.
Huan C (2016) Leaders or readers, whom to please? News values in the transition of the Chinese press. *Discourse, Context & Media* 13: 114-121.
Iedema R (2003) *Discourses of post-bureaucratic organization.* Amsterdam: John Benjamins Publishing.
Igwebuike EE and Chimuanya L (2020) Legitimating falsehood in social media: A discourse analysis of political fake news. *Discourse & Communication* 15(1): 42-58.
Inwood O and Zappavigna M (2021) Ambient affiliation, misinformation, and moral panic: Negotiating social bonds in a YouTube internet hoax. *Discourse & Communication* 15(3): 281-307.

Inwood O and Zappavigna M (2022a) The ID2020 Conspiracy Theory in YouTube Video Comments during COVID-19: Bonding around Religious, Political, and Technological Discourses. In M Demata, V Zorzi and A Zottola (eds) *Conspiracy Theory Discourses*. Amsterdam: John Benjamins, pp.241-266.

Inwood O and Zappavigna M (2022b) A Systemic Functional Linguistics Approach to Analysing White Supremacist and Conspiratorial Discourse on YouTube. *The Communication Review* 25(3-4): 204-234

Inwood O and Zappavigna M (2023) Conspiracy Theories and White Supremacy on YouTube: Exploring Affiliation and Legitimation Strategies in YouTube Comments. *Social Media + Society* 9(1): 1-16.

Inwood O and Zappavigna M (forthcoming) Attitudes about Propaganda and Disinformation: Identifying Discursive Personae in YouTube Comment Sections. In S Maci, M Demata, M McGlashan and P Seargeant (eds) *Routledge Handbook of Discourse and Disinformation*. London: Routledge.

Jane EA and Fleming C (2014) *Modern conspiracy: The importance of being paranoid.* New York: Bloomsbury Publishing.

Jaynes V (2020) The social life of screenshots: the power of visibility in teen friendship groups. *New media & society* 22(8): 1378-1393.

Jenkins BM and Cramer EM (2022) Capturing Injustice: The Screenshot as a Tool for Sousveillance. *Howard Journal of Communications*. 1-22.

Jenkins H, Ford S and Green J (2013) Spreadable media. *Spreadable Media*. New York: New York University Press.

Jeon YA, Hale B, Knackmuhs E and Mackert M (2018) Weight stigma goes viral on the Internet: Systematic assessment of YouTube comments attacking overweight men and women. *Interactive Journal of Medical Research* 7(1): 1-11.

Jepperson R and Meyer JW (2011) Multiple Levels of Analysis and the Limitations of Methodological Individualisms. *Sociological Theory* 29(1): 54-73.

Jewitt C, Bezemer J and O'Halloran K (2016) *Introducing Multimodality.* London: Routledge.

Jin F, Wang W, Zhao L, et al. (2014) Misinformation propagation in the age of twitter. *IEEE Computer* 47(12): 90-94.

Johns A (2017) Flagging White Nationalism 'After Cronulla': From the Beach to the Net. *Journal of Intercultural Studies* 38(3): 349-364.

Kalbfleisch PJ and Docan-Morgan T (2019) Defining Truthfulness, Deception, and Related Concepts. In T Docan-Morgan (Ed) *The Palgrave Handbook of Deceptive Communication*. Cham, Switzerland: Springer, pp.29-39.

Karimi N, Moore A and Lukin A (2018) Cancer care as an integrated practice: Consultations between an oncologist and patients with advanced, incurable cancer. *Perspectives from Systemic Functional Linguistics*. London: Routledge, pp.315-337.

Karlova NA and Fisher KE (2013) "Plz RT": A Social Diffusion Model of Misinformation and Disinformation for Understanding Human Information Behaviour. *Information Research* 18.1: 1-17.

Kashima Y, Dennis S, Perfors A and Laham SM (2021) Culture and global societal threats: COVID-19 as a pathogen threat to humanity. *Group Processes & Intergroup Relations* 24(2): 223-230.

Keller FB, Schoch D, Stier S and Yang J (2020) Political astroturfing on Twitter: How to coordinate a disinformation campaign. *Political Communication* 37(2): 256-280.
Klausen J, Barbieri ET, Reichlin-Melnick A and Zelin AY (2012) The YouTube Jihadists: A social network analysis of Al-Muhajiroun's propaganda campaign. *Perspectives on Terrorism* 6(1): 36-53.
Knight NK (2008) "Still cool…and american too!": an SFL analysis of deferred bonds in internet messaging humour. In N Nørgaard (ed) *Systemic Functional Linguistics in Use, Odense Working Papers in Language and Communication*. pp.481-502.
Knight NK (2010a) *Laughing our bonds off: Conversational humour in relation to affiliation*. Unpublished PhD Thesis: University of Sydney.
Knight NK (2010b) Wrinkling complexity: Concepts of identity and affiliation in humour. In M Bednarek and JR Martin (eds) *New discourse on language: Functional perspectives on multimodality, identity, and affiliation.* London/New York: Continuum, 35–58.
Knight NK (2013) Evaluating experience in funny ways: how friends bond through conversational humour. *Text & Talk* 33(4-5): 553-574.
Konkes C and Lester L (2017) Incomplete knowledge, rumour and truth seeking: when conspiracy theories become news. *Journalism Studies* 18(7): 826-844.
Kopytowska M (2021) Xenophobia, misogyny and rape culture: Targeting women in cyberspace. *Journal of Language Aggression and Conflict* 9(1): 76-99.
Koteyko N and Hunt D (2016) Performing health identities on social media: An online observation of Facebook profiles. *Discourse, Context & Media* 12: 59-67.
Krafft P and Donovan J (2020) Disinformation by Design: The Use of Evidence Collages and Platform Filtering in a Media Manipulation Campaign. *Political Communication* 37(2): 194-214.
Kress GR and Van Leeuwen T (2006) *Reading images: The grammar of visual design (Second Edition).* London: Routledge.
Krzyżanowski M (2018) Discursive shifts in ethno-nationalist politics: On politicization and mediatization of the "refugee crisis" in Poland. *Journal of Immigrant & Refugee Studies* 16(1-2): 76-96.
Krzyżanowski M, Triandafyllidou A and Wodak R (2018) The mediatization and the politicization of the "refugee crisis" in Europe. *Journal of Immigrant & Refugee Studies* 16(1-2): 1-14.
Kumar KPK and Geethakumari G (2014) Detecting misinformation in online social networks using cognitive psychology. *Human-centric Computing and Information Sciences* 4(1): 14.
Kumar S, West R and Leskovec J (2016) Disinformation on the web: Impact, characteristics, and detection of wikipedia hoaxes. *Proceedings of the 25th International Conference on World Wide Web:* 591-602.
Labov W (1972) *Sociolinguistic patterns.* Pennsylvania: University of Pennsylvania Press.
Lazer DM, Baum MA, Benkler Y, Berinsky A, Greenhill K, Menczer F, Metzger M, Nyhan B, Pennycock G, Roshschild D, Schudson M, Sloman S, Sunstein C, Thorson E, Watts D and Zittrain J (2018) The science of fake news. *Science* 359(6380): 1094-1096.
Levine TR (2014) Truth-default theory (TDT) a theory of human deception and deception detection. *Journal of Language and Social Psychology* 33(4): 378-392.

Levy H (2020) Grammars of contestation and pluralism: Paulo Freire's action in Brazil's periphery and the rise of right-wing discourse on YouTube. *International Communication Gazette* 82(5): 474-489.
Lewis R (2018) Alternative influence: Broadcasting the reactionary right on YouTube. *New Data & Society [ebook]. Retrieved from: https://datasociety.net/output/alternative-influence*. Accessed 1st Feb 2023.
Lewis SP, Heath NL, Sornberger MJ, et al. (2012) Helpful or harmful? An examination of viewers' responses to nonsuicidal self-injury videos on YouTube. *Journal of Adolescent Health* 51(4): 380-385.
Liebmann B (2018) *How Long Does Content Last and How Frequently Should You Post on Social Media?* Available at: https://sprocketwebsites.com/Blog/how-long-does-content-last-and-how-frequently-should-you-post-on-social-media. Accesed 1st Feb 2023.
Lim SS (2020) Manufacturing Hate 4.0: Can Media Studies Rise to the Challenge? *Television & New Media* 21(6): 602-607.
Liu F and Hood S (2019) Rhetorical strategies of political persuasion: The play of irrealis and realis meaning in re/aligning readers in newspaper editorials. *Text & Talk* 39(5): 589-611.
Logi L and Zappavigna M (2021) Impersonated personae–paralanguage, dialogism and affiliation in stand-up comedy. *HUMOR* 34(3) 339-373.
Lucas E (2014) *The new Cold War: Putin's Russia and the threat to the West.* New York: Macmillan.
Lukasik M, Cohn T and Bontcheva K (2015) Classifying tweet level judgements of rumours in social media. *arXiv preprint arXiv:1506.00468*.
Mackay RR (2015) Multimodal legitimation: selling Scottish independence. *Discourse & Society* 26(3): 323-348.
Mahl D, Schäfer MS and Zeng J (2022) Conspiracy theories in online environments: An interdisciplinary literature review and agenda for future research. *New Media & Society*. OnlineFirst: 1-21.
Makki M (2019) 'Discursive news values analysis' of Iranian crime news reports: Perspectives from the culture. *Discourse & Communication* 13(4): 437-460.
Makki M (2020) The role of 'culture'in the construction of news values: a discourse analysis of Iranian hard news reports. *Journal of Multicultural Discourses* 15(3): 308-324.
Malinowski B (2001) *Coral Gardens and Their Magic: A Study of the Methods of Tilling the Soil and of Agricultural Rites in the Trobriand Islands. The language of magic and gardening. II.* London: Psychology Press.
Marshall PD, Moore C and Barbour K (2015) Persona as method: exploring celebrity and the public self through persona studies. *Celebrity Studies* 6(3): 288-305.
Marshall PD, Moore C and Barbour K (2019) *Persona studies: an introduction.* New Jersey: John Wiley & Sons.
Martin JR (1989) *Factual writing: Exploring and challenging social reality.* Oxford: Oxford University Press.
Martin JR (1992) *English text: System and structure.* Amsterdam: John Benjamins Publishing.
Martin JR (2008a) Innocence: realisation, instantiation and individuation in a Botswanan town. *Questioning linguistics*. 32-76.

Martin JR (2008b) Tenderness: Realisation and instantiation in a Botswanan town. *Odense working papers in language and communication* 29: 30-58.
Martin JR (2009) Realisation, instantiation and individuation: some thoughts on identity in youth justice conferencing. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada* 25(spe): 549-583.
Martin JR (2016) Meaning matters: A short history of systemic functional linguistics. *Word* 62(1): 35-58.
Martin JR and Rose D (2003) *Working with discourse: Meaning beyond the clause.* London: Bloomsbury Publishing.
Martin JR and Rose D (2008) *Genre relations: Mapping culture.* London: Equinox.
Martin JR and Stenglin M (2007) Materializing reconciliation: Negotiating difference in a transcolonial exhibition. In T Royce and W Bowcher (Eds) *New Directions in the Analysis of Multimodal Discourse*. 215-238.
Martin JR and White PRR (2005) *The Language of Evaluation: Appraisal in English.* New York: Palgrave Macmillan.
Martin JR and Wodak R (2003) *Re/reading the past: Critical and functional perspectives on time and value.* Amsterdam: John Benjamins Publishing.
Martinec R (2004) Gestures that co-occur with speech as a systematic resource: the realization of experiential meanings in indexes. *Social Semiotics* 14(2): 193-213.
Matamoros-Fernández A (2017) Platformed racism: the mediation and circulation of an Australian race-based controversy on Twitter, Facebook and YouTube. *Information, Communication & Society* 20(6): 930-946.
Matamoros-Fernández A and Farkas J (2021) Racism, Hate Speech, and Social Media: A Systematic Review and Critique. *Television & New Media* 22(2): 205-224.
Maton K, Hood S and Shay S (2015) *Knowledge-building: educational studies in legitimation code theory.* London: Routledge.
Matthiessen CM (2013) Applying systemic functional linguistics in healthcare contexts. *Text & Talk* 33(4-5): 437-466.
Matthiessen CM, Teruya K and Lam M (2010) *Key Terms in Systemic Functional Linguistics.* London: A&C Black.
McCornack SA (1992) Information manipulation theory. *Communication Monographs* 59(1): 1-16.
McDonald D and Woodward-Kron R (2016) Member roles and identities in online support groups: Perspectives from corpus and systemic functional linguistics. *Discourse & Communication* 10(2): 157-175.
McDougall J (2019) *Fake News vs Media Studies: Travels in a False Binary.* Cham: Palgrave Macmillan.
McGlone MS and Knapp ML (2019) Historical Perspectives on the Study of Lying and Deception. In T Docan-Moragan (ed) *The Palgrave Handbook of Deceptive Communication*. Cham: Springer, pp.3-28.
McNair B (2017) *Fake news: Falsehood, fabrication and fantasy in journalism.* New York: Routledge.
Mejias UA and Vokuev NE (2017) Disinformation and the media: the case of Russia and Ukraine. *Media, Culture & Society* 39(7): 1027-1042.
Mendoza M, Poblete B and Castillo C (2010) Twitter Under Crisis: Can we trust what we RT? *Proceedings of the First Workshop on Social Media Analytics.* ACM, 71-79.

Metts S (1989) An exploratory investigation of deception in close relationships. *Journal of Social and Personal Relationships* 6(2): 159-179.
Miller DT (2021) Characterizing QAnon: Analysis of YouTube comments presents new conclusions about a popular conservative conspiracy. *First Monday* 26(2).
Mintz A (2012) *Web of Deceit: Misinformation and Manipulation in the Age of Social Media.* Medford: CyberAge Books.
Mohammed SN (2019) Conspiracy Theories and Flat-Earth Videos on YouTube. *The Journal of Social Media in Society* 8(2): 84-102.
Molek-Kozakowska K (2017) Communicating environmental science beyond academia: Stylistic patterns of newsworthiness in popular science journalism. *Discourse & Communication* 11(1): 69-88.
Moreno López G (2020) A multimodal discourse analysis of the most viewed reggaeton video on youtube by the LIV super bowl halftime show. *Open Journal for Studies in Arts* 3(2).
Murphy B (2022) Decaying National Security and the Rise of Imagined Tribalism. *The RUSI Journal*. 1-13.
Murthy D and Sharma S (2019) Visualizing YouTube's comment space: online hostility as a networked phenomena. *New Media & Society* 21(1): 191-213.
Näsänen J (2017) Two versions of nomadic employees: Opposing ways to employ the same discourse in talking about change. *Discourse & Communication* 11(3): 259-275.
Nohrstedt SA (2013) Mediatization as an echo-chamber for xenophobic discourses in the threat society: The Muhammad cartoons in Denmark and Sweden. In R Wodak, M Khosravinok and B Mral (eds) *Right-Wing Populism in Europe. Politics and Discourse*. London: Bloomsbury, pp. 309-321.
O'Halloran KL (2008) Systemic functional-multimodal discourse analysis (SF-MDA): Constructing ideational meaning using language and visual imagery. *Visual Communication* 7(4): 443-475.
O'Halloran KL (2004) Visual semiosis in film. In KL O'Halloran (ed) *Multimodal discourse analysis: Systemic functional perspectives*. London: Continuum International Publishing, pp. 109-130.
O'Halloran KL, Tan S, Pham D-S, et al. (2018) A Digital Mixed Methods Research Design: Integrating Multimodal Analysis With Data Mining and Information Visualization for Big Data Analytics. *Journal of Mixed Methods Research* 12(1): 11-30.
O'Halloran KL, Tan S, Wignell P, et al. (2019) Interpreting text and image relations in violent extremist discourse: A mixed methods approach for big data analytics. *Terrorism and Political Violence* 31(3): 454-474.
O'Toole M (1994) *The Language of Displayed Art.* New Jersey: Fairleigh Dickinson University Press.
O'Toole M (2004) Opera Ludentes: the Sydney Opera House at work and play. In KL O'Halloran (ed) *Multimodal discourse analysis: Systemic functional perspectives*. London: Bloomsbury Publishing, pp. 11-27.
Oddo J (2011) War legitimation discourse: Representing 'Us' and 'Them'in four US presidential addresses. *Discourse & Society* 22(3): 287-314.
Page RE (2003) An analysis of APPRAISAL in childbirth narratives with special consideration of gender and storytelling style. *Text & Talk* 23(2): 211-237.

Painter C (1996) The development of language as a resource for thinking: A linguistic view of learning. In R Hasan and G Williams (eds) *Literacy in Society*. London: Longman, pp. 50-85.

Panizo-LLedot A, Torregrosa J, Bello-Orgaz G, et al. (2019) Describing alt-right communities and their discourse on twitter during the 2018 us mid-term elections. *International Conference on Complex Networks and their Applications.* Cham: Springer, pp. 427-439.

Paolillo JC (2018) The Flat Earth phenomenon on YouTube. *First Monday* 23(12).

Patrick J (2006) The Scamseek Project – Text Mining for Financial Scams on the Internet. In GJ Williams and SJ Simoff (eds) *Data Mining: Lecture Notes in Computer Science, vol 3755,* Cham: Springer, pp. 295-302.

Patterson AN (2018) YouTube Generated Video Clips as Qualitative Research Data: One Researcher's Reflections on the Process. *Qualitative Inquiry* 24(10): 759-767.

Peer L and Ksiazek TB (2011) YouTube and The Challenge to Journalism. *Journalism Studies* 12(1): 45-63.

Peled-Elhanan N (2010) Legitimation of massacres in Israeli school history books. *Discourse & Society* 21(4): 377-404.

Pérez-Arredondo C and Cárdenas-Neira C (2019) Space and legitimation: The multimodal representation of public space in news broadcast reports on Hooded Rioters. *Discourse & Communication* 13(3): 279-302.

Pigden CR (2007) Conspiracy theories and the conventional wisdom. *Episteme: A Journal of Social Epistemology* 4(2): 219-232.

Pihlaja S (2014) *Antagonism on YouTube: Metaphor in online discourse.* London: Bloomsbury Publishing.

Polyakova A and Boyer SP (2018) The future of political warfare: Russia, the West, and the coming age of global digital competition. *Foreign Policy at Brookings, March 2018: www.brookings.edu/wp-content/uploads/2018/03/the-future-of-political-warfare.pdf* Accessed 1st Feb 2023.

Potter J and Wetherell M (2002) Analyzing discourse. In A Bryman and RG Burgess (Eds) *Analyzing qualitative data*. London: Routledge, pp.61-80.

Poynton C (1984) Names as vocatives: Forms and functions. *Nottingham Linguistic Circular* 13: 1-34.

Poynton C (1996) Amplification as a grammatical prosody: attitudinal modification in the nominal group. *Advances in Discourse Processes* 57: 211-228.

Procter R, Vis F and Voss A (2013) Reading the riots on Twitter: methodological innovation for the analysis of big data. *International Journal of Social Research Methodology* 16(3): 197-214.

Qazvinian V, Rosengren E, Radev DR, et al. (2011) Rumor has it: Identifying misinformation in microblogs. *Proceedings of the Conference on Empirical Methods in Natural Language Processing.* Association for Computational Linguistics, 1589-1599.

Qisthi A, Ricahyono R and Chasanatun T (2022) Contexts of video on devina hermawan's youtube channel: A functional discourse approach. *Social Sciences, Humanities and Education Journal (SHE Journal)* 3(1): 52-64.

Ravelli LJ (2000) Beyond shopping: Constructing the Sydney Olympics in three-dimensional text. *Text & Talk* 20(4): 489-516.

Ravelli LJ and McMurtrie RJ (2015) *Multimodality in the built environment: Spatial discourse analysis.* London: Routledge.
Ravelli LJ and Van Leeuwen T (2018) Modality in the digital age. *Visual Communication* 17(3): 277-297.
Reilly I (2012) Satirical fake news and/as American political discourse. *The Journal of American Culture* 35(3): 258-275.
Restrepo NJ, Illari L, Leahy R, Sear RF, Lupu Y and Johnson NF (2021) How Social Media Machinery Pulled Mainstream Parenting Communities Closer to Extremes and Their Misinformation During Covid-19. In *IEEE Access* 10: 2330-2344.
Richardson JE and Wodak R (2009) The Impact of Visual Racism: Visual Arguments in Political Leaflets of Austrian and British Far-right Parties. *Controversia* 6(2): 45-77.
Rieder B (2015) YouTube Data Tools (Version 1.11). [Software]. Available from: https://tools.digitalmethods.net/netvizz/youtube/. Accessed 1st Feb 2023.
Rieder B, Matamoros-Fernández A and Coromina Ò (2018) From ranking algorithms to 'ranking cultures' Investigating the modulation of visibility in YouTube search results. *Convergence* 24(1): 50-68.
Rivers DJ and Ross AS (2020) Authority (de) legitimation in the border wall Twitter discourse of President Trump. *Journal of Language and Politics* 19(5): 831-856.
Rizwan S (2019) Legitimation strategies and theistic worldview in sociopolitical discourse: A systemic functional critical discourse analysis of Pakistani social media discussions. *Text & Talk* 39(2): 235-260.
Robins RH (2004) Malinowski, Firth, and the'Context of Situation'. In E Ardener (Ed) *Social Anthropology and Language*. London: Tavistock, 33-46.
Robles J (2017) Exclusion in gossipy talk: Highjacking the preference structure for ingroup belonging. *CADAAD Journal* 9(2): 5-22.
Röchert D, Shahi GK, Neubaum G, Ross B and Stieglitz S (2021) The Networked Context of COVID-19 Misinformation: Informational Homogeneity on YouTube at the Beginning of the Pandemic. *Online Social Networks and Media* 26: 1-13.
Rosnow RL and Foster EK (2005) Rumor and gossip research. *Psychological Science Agenda* 19(4): 1-2.
Ross AS (2020) Discursive delegitimisation in metaphorical# secondcivilwarletters: an analysis of a collective Twitter hashtag response. *Critical Discourse Studies* 17(5): 510-526.
Ross AS and Rivers DJ (2017) Digital cultures of political participation: Internet memes and the discursive delegitimization of the 2016 US Presidential candidates. *Discourse, Context & Media* 16: 1-11.
Rubin V, Conroy N, Chen Y and Cornwell S. (2016) Fake news or truth? using satirical cues to detect potentially misleading news. *Proceedings of the Second Workshop on Computational Approaches to Deception Detection.* 7-17.
Sellnow TL, Parrish A and Semenas L (2019) From Hoax as Crisis to Crisis as Hoax: Fake News and Information Disorder as Disruptions to the Discourse of Renewal. *Journal of International Crisis and Risk Communication Research* 2(1): 6.
Senft TM (2013) Microcelebrity and the branded self. In J Hartley, J Burgess and A Bruns (eds) *A Companion to New Media Dynamics*. West Sussex: John Wiley & Sons, pp.346-354.

Shahsavari S, Holur P, Wang T, et al. (2020) Conspiracy in the time of corona: Automatic detection of emerging COVID-19 conspiracy theories in social media and the news. *Journal of Computational Social Science* 3(2): 279-317.
Shao C, Ciampaglia GL, Flammini A and Menczer F (2016) Hoaxy: A Platform for Tracking Online Misinformation. *The 25th International Conference Companion.* ACM Press, 745-750.
Shao C, Ciampaglia GL, Varol O, Flammini A and Menczer F (2017) The spread of fake news by social bots. *arXiv preprint arXiv:1707.07592*.
Shooman Y (2016) Between Everyday Racism and Conspiracy Theories: Islamophobia on the German-Language Internet. In G Ruhrmann, Y Shooman and P Widmann, *Media and Minorities: Questions on Representation from an International Perspective.* Berlin: Vandenhoeck & Ruprecht, pp. 136-154.
Sindoni MG (2020) '# YouCanTalk': A multimodal discourse analysis of suicide prevention and peer support in the Australian BeyondBlue platform. *Discourse & Communication* 14(2): 202-221.
Singh VK, Ghosh I and Sonagara D (2021) Detecting fake news stories via multimodal analysis. *Journal of the Association for Information Science and Technology* 72(1): 3-17.
Smith CA (2019) Weaponized iconoclasm in Internet memes featuring the expression 'Fake News'. *Discourse & Communication* 13(3): 303-319.
Smith LG, Blackwood L and Thomas EF (2020) The need to refocus on the group as the site of radicalization. *Perspectives on Psychological Science* 15(2): 327-352.
Søe SO (2018) Algorithmic detection of misinformation and disinformation: Gricean perspectives. *Journal of Documentation* 74(2): 309-332.
Starbird K, Arif A and Wilson T (2019) Disinformation as collaborative work: Surfacing the participatory nature of strategic information operations. *Proceedings of the ACM on Human-Computer Interaction* 3(CSCW): 1-26.
Stenglin M (2008) Interpersonal meaning in 3D space: How a bonding icon gets its 'charge'. In L Unsworth (Ed) *Multimodal semiotics: Functional analysis in contexts of education*. London: Bloomsbury, pp. 50-66.
Stenglin M and Cléirigh C (2020) Scientific grant application writing: Re/packaging text to enhance its impact. *Journal of English for Academic Purposes* 44: 1-18.
Stenglin MK (2004) Packaging curiosities: Towards a grammar of three-dimensional space. Unpublished PhD Thesis, University of Sydney.
Stern B (1994) Authenticity and the textual persona: postmodern paradoxes in advertising narrative. *International Journal of Research in Marketing* 11(4): 387-400.
Strauss A and Corbin JM (1997) *Grounded theory in practice.* Thousand Oaks: Sage.
Sundström M and Obenius H (2020) (De-) legitimation of migration: A critical study of social media discourses. In D Abdelhady, N Gren and M Joormann (Eds) *Refugees and the violence of welfare bureaucracies in Northern Europe*. Manchester: Manchester University Press, pp. 144-158.
Suriyanti A (2022) Texts, Images, And Music in Billie Eilish's Bad Guy Music Video on Youtube. Com. *International Journal of Systemic Functional Linguistics* 5(1): 50-56.

Szenes E (2017) The linguistic construction of business reasoning: Towards a language-based model of decision-making in undergraduate business. Unpublished PhD Thesis, University of Sydney.
Szenes E (2021) Neo-Nazi environmentalism: The linguistic construction of ecofascism in a Nordic Resistance Movement manifesto. *Journal for Deradicalization*.(27): 146-192.
Tandoc Jr EC, Lim ZW and Ling R (2018) Defining "fake news" A typology of scholarly definitions. *Digital Journalism* 6(2): 137-153.
Tann K (2011) Semogenesis of a Nation. Unpublished PhD Thesis, University of Sydney.
Tann K (2012) The language of identity discourse: introducing a systemic functional framework for iconography. *Linguistics & the Human Sciences* 8(3): 361-391.
Taverniers M (2011) The syntax–semantics interface in systemic functional grammar: Halliday's interpretation of the Hjelmslevian model of stratification. *Journal of Pragmatics* 43(4): 1100-1126.
Teich E (1999) *Systemic Functional Grammar & Natural Language Generation.* London: A&C Black.
Thelwall M (2018) Social media analytics for YouTube comments: Potential and limitations. *International Journal of Social Research Methodology* 21(3): 303-316.
Thelwall M, Sud P and Vis F (2012) Commenting on YouTube videos: From Guatemalan rock to el big bang. *Journal of the American Society for Information Science and Technology* 63(3): 616-629.
Tian P (2010) Playing with "Femininity": An Intermodal Analysis of the Bilingual Picture Book The Ballad of Mulan. In JR Marin and M Bednarek (eds) *New Discourse on Language: Functional Perspectives on Multimodality, Affiliation*. London: Bloomsbury Publishing, p. 134-162.
Turner RE, Edgley C and Olmstead G (1975) Information control in conversations: Honesty is not always the best policy. *Kansas Journal of Sociology*: 11(1): 69-89.
Unsworth L (2005) *Researching language in schools and communities: Functional linguistic perspectives.* London: A&C Black.
Vachek J (1964) *A Prague School reader in linguistics.* Indiana: Indiana University Press.
Vachek J (1966) *The Linguistic School of Prague: An introduction to its theory and practice.* Indiana: Indiana University Press.
Van der Linden S and Roozenbeek J (2020) Psychological inoculation against fake news. In RGreifeneder, ME Jaffe, EJ Newman and N Schwarz (eds) *The Psychology of Fake News: Accepting, sharing, and correcting misinformation*. Milton Park: Routledge, pp.147-169.
Van Dijk TA (1997) *Discourse as structure and process.* Thousand Oaks: Sage.
Van Dijk TA (1998) *Ideology: A multidisciplinary approach.* Thousand Oaks: Sage.
van Dijk TA (2006) Discourse and manipulation. *Discourse and Society* 17 (3): 359-383.
Van Leeuwen T (1991) The sociosemiotics of easy listening music. *Social Semiotics* 1(1): 67-80.
Van Leeuwen T (1995) The grammar of legitimation. *London: School of Printing, School of Media*.
Van Leeuwen T (1999) *Speech, music, sound.* London: Bloomsbury Publishing.
Van Leeuwen T (2005) *Introducing social semiotics.* London: Psychology Press.
Van Leeuwen T (2007) Legitimation in discourse and communication. *Discourse & Communication* 1(1): 91-112.

Van Leeuwen T (2008) *Discourse and practice: New tools for critical discourse analysis.* Oxford: Oxford University Press.
Van Leeuwen T (2021) *Multimodality and Identity.* London: Routledge.
Van Leeuwen T and Wodak R (1999) Legitimizing immigration control: A discourse-historical analysis. *Discourse studies* 1(1): 83-118.
Van Prooijen J-W, Spadaro G and Wang H (2022) Suspicion of institutions: How distrust and conspiracy theories deteriorate social relationships. *Current Opinion in Psychology* 43: 65-69.
Vanek J, King K and Bigelow M (2018) Social presence and identity: Facebook in an English language classroom. *Journal of Language, Identity & Education* 17(4): 236-254.
Vásquez C (2014) 'Usually not one to complain but…': Constructing identities in user-generated online reviews. In P Seargeant and C Tagg (eds) *The Language of Social Media*. Cham: Springer, pp.65-90.
Vincent JM and Castelfranchi C (1979) On the art of deception: How to lie while saying the truth. In H Parret, M Sbisa and J Verschueren (Eds) *Conference on Pragmatics*, Trente, IT: Urbino.
Vis F (2013) Twitter as a reporting tool for breaking news: Journalists tweeting the 2011 UK riots. *Digital journalism* 1(1): 27-47.
Volkova S, Shaffer K, Jang JY and Hodas N. (2017) Separating Facts from Fiction: Linguistic Models to Classify Suspicious and Trusted News Posts on Twitter. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers).* Association for Computational Linguistics, 647-653.
Vraga EK, Kim SC and Cook J (2019) Testing Logic-based and Humor-based Corrections for Science, Health, and Political Misinformation on Social Media. *Journal of Broadcasting & Electronic Media* 63(3): 393-414.
Wallace LE, Wegener DT and Petty RE (2020) When sources honestly provide their biased opinion: Bias as a distinct source perception with independent effects on credibility and persuasion. *Personality and Social Psychology Bulletin* 46(3): 439-453.
Wardle C and Derakhshan H (2017) Information Disorder: Toward an interdisciplinary framework for research and policy making. *Council of Europe report, DGI(2017)* 9. www.coe.int Accessed 1st Feb 2023.
Wardle C and Singerman E (2021) Too little, too late: social media companies' failure to tackle vaccine misinformation poses a real threat. *BMJ (Clinical Research)* 372.
Wasserman H (2017) Fake news from Africa: Panics, politics and paradigms. *Journalism: Theory, Practice & Criticism*. 21(1): 3-16.
White PRR (1998) *Telling media tales: The news story as rhetoric.* Unpublished PhD Thesis, University of Sydney.
White PRR (2008) Praising and blaming, applauding, and disparaging–solidarity, audience positioning, and the linguistics of evaluative disposition. In G Antos and E Ventola (Eds) *Handbook of Interpersonal Communication*. Berlin: De Gruyter Mouton, pp.567-594.
White PRR and Sano M (2006) Dialogistic positions and anticipated audiences-a framework for stylistic comparisons. In K Aijmer and A Simon-Vandenbergen (Eds) *Pragmatic Markers in Contrast*. Frankfurt: Elsevier, pp.189-214.

Wignell P, Chai K, Tan S, et al. (2021) Natural language understanding and multimodal discourse analysis for interpreting extremist communications and the re-use of these materials online. *Terrorism and political violence* 33(1): 71-95.
Wignell P, Tan S and O'Halloran KL (2017a) Under the shade of AK47s: A multimodal approach to violent extremist recruitment strategies for foreign fighters. *Critical Studies on Terrorism* 10(3): 429-452.
Wignell P, Tan S, O'Halloran KL, et al. (2017b) A mixed methods empirical examination of changes in emphasis and style in the extremist magazines Dabiq and Rumiyah. *Perspectives on Terrorism* 11(2): 2-20.
Wodak R (2015) *The politics of fear: What right-wing populist discourses mean.* Thousand Oaks: Sage.
Yablokov I (2015) Conspiracy Theories as a Russian Public Diplomacy Tool: The Case of Russia Today (RT). *Politics* 35(3-4): 301-315.
Yarow J and Angelova K (2010) *Chart of the Day - The Lifecycle of a YouTube Video* Available at: https://www.businessinsider.com.au/chart-of-the-day-the-lifecycle-of-a-youtube-video-2010-5?r=US&IR=T. Accessed 1st Feb 2023.
Zappavigna M (2011) Ambient affiliation: A linguistic perspective on Twitter. *New Media & Society* 13(5): 788-806.
Zappavigna M (2012) *The Discourse of Twitter and Social Media: How we use language to create affilation on the web.* London/New York: Continuum.
Zappavigna M (2014a) CoffeeTweets: Bonding around the bean on Twitter. In: P Seargeant and C Tagg (eds) *The Language of Social Media*. Berlin: Springer, pp.139-160.
Zappavigna M (2014b) Enacting Identity in Microblogging through Ambient Affiliation. *Discourse & Communication* 8(2): 209-228.
Zappavigna M (2018) Searchable Talk and Social Media Metadiscourse. London: Bloomsbury Publishing.
Zappavigna M (2019) The Organised Self and Lifestyle Minimalism: Multimodal Deixis and Point of View in Decluttering Vlogs on YouTube. *Multimodal Communication* 8(1): 1-14.
Zappavigna M (2021) Ambient Affiliation in Comments on YouTube Videos: Communing Around Values About ASMR. Journal of Foreign Languages *外国语* 44(1): 21-40.
Zappavigna M, Cleirigh C, Dwyer P, et al. (2010) The coupling of gesture and phonology. In JR Martin and M Bednarek (Eds) *New discourse on language: Functional perspectives on multimodality, identity, and affiliation*. New York: Bloomsbury, 219-236.
Zappavigna M and Martin JR (2017) *Discourse and diversionary justice: an analysis of youth justice conferencing.* Cham: Springer.
Zappavigna M and Martin JR (2018) #Communing affiliation: Social tagging as a resource for aligning around values in social media. *Discourse, Context & Media (Special issue on the discourse of social tagging)* 22: 1-64.
Zappavigna M and Zhao S (2017) Selfies in 'mommyblogging': An emerging visual genre. *Discourse, Context & Media* 20: 239-247.
Zerback T, Töpfl F and Knöpfle M (2021) The disconcerting potential of online disinformation: Persuasive effects of astroturfing comments and three strategies for inoculation against them. *New Media & Society* 23(5): 1080-1098.

Zhao S and Zappavigna M (2018a) Beyond the self: Intersubjectivity and the social semiotic interpretation of the selfie. *New Media & Society* 20(5): 1735-1754.
Zhao S and Zappavigna M (2018b) The interplay of (semiotic) technologies and genre: The case of the selfie. *Social Semiotics* 28(5): 665-682.
Zhou L and Zhang D (2007) An ontology-supported misinformation model: Toward a digital misinformation library. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 37(5): 804-813.
Zubiaga A, Liakata M and Procter R (2017) Exploiting Context for Rumour Detection in Social Media. In G Ciampaglia, A Mashhadi and T Yasseri (eds) *International Conference on Social Informatics*. Cham: Springer, pp. 109-123.
Zubiaga A, Liakata M, Procter R, Bontcheva K and Tolmie P. (2015) Towards detecting rumours in social media. *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*.
Zuk P and Zuk P (2020) Right-wing populism in Poland and anti-vaccine myths on YouTube: Political and cultural threats to public health. *Global Public Health* 15(6): 790-804.