UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING AND PHYSICAL SCIENCES

Electronics and Computer Science

Modelling Eristic and Rhetorical Argumentation on the Social Web

by

Tom Blount

Thesis for the degree of Doctor of Philosophy

September 2018

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND PHYSICAL SCIENCES
Electronics and Computer Science

Doctor of Philosophy

MODELLING ERISTIC AND RHETORICAL ARGUMENTATION ON THE SOCIAL WEB

by Tom Blount

Argumentation, debate and discussion are key facets of human communication, shaping the way people form, share and promote ideas, hypotheses and solutions to problems. Argumentation can broadly be broken down into collaborative problem solving or truth-seeking (dialectic argumentation) and quarrelling without hope for a resolution, either aggressively or for the purpose of recreation, catharsis or entertainment (eristic argumentation). Techniques used within argumentation can likewise be classified as primarily fact-based (logical), or emotion-based (rhetorical).

The social web, consisting of the people, tools and communities that form over the world wide web, is a growing way in which individuals, social groups and even corporations share content, ideas and information, as well as hold discussions and debates. Current models of argumentation often focus on formal argumentation techniques, in which participants are expected to abide by a stringent set of rules or practices. However, on the social web there is no such code of conduct. Antisocial behaviour, which often stems from argumentation, can have a negative impact on online communities, driving away new users and stifling participation.

This thesis examines the way in which the use of eristic and rhetorical argumentation impacts the perception and engagement of participants in, and the audience of, arguments on the social web. After a preliminary investigation to determine the effectiveness with which current formal models represent eristic argument, a series of augmentations to these existing models was proposed, dubbed the Argumentation on the Social Web Ontology. This allows for the explicit representation of rhetorical support and attack, and supports the annotation of groups of participants and the viewing audience. These augmentations were further refined through deeper investigation of modelling social argument, and through expert review. This culminated in the creation of a large dataset which was in turn used to drive an experiment into the way in which social media users perceive and engage with different types of eristic argumentation, showing that while rhetorical tactics were often more entertaining and offensive than their logical counterparts, they did not significantly alter the degree of engagement with the discussion.

Contents

List of Figures

List of Tables

Declaration of Authorship

I, Tom Blount , declare that the thesis entitled *Modelling Eristic and Rhetorical Argumentation on the Social Web* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as: (Blount et al., 2014), (Blount et al., 2015a), (Blount et al., 2015b) and (Blount et al., 2016)

Signed:.....

Date:.....

Acknowledgements

First and foremost, thanks must go to my supervisors, David Millard and Mark Weal, for their ongoing advice and guidance, and without who this thesis would not be possible, as well as my examiners Nicholas Gibbins and Floris Bex, whose feedback shaped the final thesis.

Equally, this would not have been possible without the love and support of my girlfriend Natasha and my family: my Dad, John; my Mum, Lin; and my Sister, Esther; all of whom have constantly been a source of encouragement and understanding, and have never stopped being there for me.

My friends Charlie Hargood, Rikki Prince, Jonathan Scott have also all provided invaluable support in the process of putting together this thesis in the form of sage academic advice, invaluable encouragement, and—most importantly of all—a welcome distraction when needed.

In addition, thanks many go to the participants of the experiments and expert review; without their contribution this work would not have been able to progress to the stage it has done.

Finally, my thanks go to the research community at large; every paper written (and reviewed), every conference attended, and every passing discussion has helped shape this thesis.

Chapter 1

Introduction

"A man may be objectively in the right, and nevertheless in the eyes of by-standers, and sometimes in his own, he may come off worst"—Schopenhauer, The Art of Always Being Right

1.1 Problem Space and Motivation

Argumentation is fundamental to human communication – it is how people share new information and new ideas, and propose new courses of action (Hahn et al., 2005; Moor and Aakhus, 2006). As a result, there is a large amount of research on argumentation from a wide variety of disciplines and topics, including: philosophy, and the nature of fallacies and how they may be critically appraised (Tindale, 2007); sociology, and the need to differentiate between classical logic and social argumentation due to the need for the capability to reason using only partial knowledge (Pólos and Hannan, 2002); law, and the need for measures of certainty and belief when modelling and reasoning over assertions (Bertea, 2004); and artificial intelligence, and the use of agent-based systems such as dialogue games, as methods for reasoning over argument to determine the victor or the correct course of action (Bench-Capon and Dunne, 2007; Karunatillake et al., 2008).

Argumentation can be (broadly) separated into two categories based on the goals and intended outcome. Firstly, dialectic argument, in which the participants are engaged in rational discourse with the aim of either discovering the particular truth behind a matter, or formulating a solution or resolution for a set of circumstances (Kerferd, 1981). Secondly, eristic argument, in which there is no clear goal and the participants are not trying to come to a resolution but are quarrelling with the aim of being seen to win, either in the eyes of their opponent or, more often, in the eyes of spectators (Kerferd, 1981; Jørgensen, 1998). Arguments can shift between these two forms, or contain "pockets" of one form within the other. Orthogonally to this, there are the notions of logic and rhetoric. While often used in modern parlance as a pejorative term, rhetoric is simply the art of discourse, and convincing an audience to one's point of view based on one's knowledge

of the topic at hand and, crucially, one's knowledge of the audience themselves (which clearly lends itself to the eristic form) whereas logic deals with reasoning between established facts (which lends itself to the dialectic form).

However, as Van Eemeren and Grootendorst (2004) note, *"perhaps out of fear of metaphysics or of 'psychologizing,' present-day logicians tend to concentrate exclusively on formalized arguments that lack any direct relation with how argumentation is conducted in practice."* Social argumentation, or the way people argue day-to-day, often has a very different structure to formalised models. In these instances, the aim of a proponent is not to prove themselves right through irrefutable logic, but simply to make others believe that they have proved themselves right.

This is particularly relevant when applied to the social web. As a network of social relationships that are created, formed and maintained through the world wide web, the social web (and the social media presented across it) are rife with discussion, debate, and argumentation (Rowe et al., 2011). As the web (and in particular the number of people, tools and communities that make up the social web) grows and becomes ubiquitous (Smith, 2009, p. 559), the potential for using it to investigate how truly massive communities interact, communicate and argue increases dramatically. However, the social web presents a number of challenges for extracting and analysing arguments, such as the lack of clear indicators of argument structure. This problem is compounded by the type of language used, which is often highly informal, incorporating slang, and irregular punctuation and grammar (Schneider et al., 2012). Further challenge includes the sheer number of distinct social platforms, which each have their own technical constraints, and cultures developed by their users (Hanna et al., 2011).

There are also a number of challenges when considering maintaining the social web as an inclusive platform for diverse and vibrant content, especially debate and discussion. There is a tendency for users to interact and associate with others who are similar in terms of traits, (such as race, age, or education) and beliefs (such as religion or politics), known as homophily (Sherchan et al., 2013) and is compounded by the introduction of "filter bubbles", the effect of content providers tailoring search results or default displays towards the preferences of individual users (Pariser, 2011). This can lead to sites becoming "echo chambers" in which well-known views and opinions are repeated, little original content is produced and there is virtually no dissent or debate (Gilbert et al., 2009). This can be further exacerbated by reputation systems, enforcing which views are acceptable in a given community by rewarding users who agree and punishing those who disagree, or those considered "outsiders" of the accepted group or culture. At the opposite end of the spectrum, where there is constant and stimulated debate, there is equal (if not greater) potential for conflict. While critical and reasonable debate, and even (respectful) recreational quarrels, are things to be encouraged, there is a visible tendency to "shout down" the opposition, including attempts to silence dissenting opinions through abuse and threats. As a result online communities can become hostile spaces, culminating in anti-social behaviour, including vulgar abuse and, at the most extreme, threats of sexual violence, and death threats (Willard, 2007; Jane, 2014).

In this thesis the case is made that disregarding social (and anti-social) interactions from argumentation models is a mistake. Accurately modelling them is the first step towards understanding exactly how argumentation is applied across the social web, and the ways in which creators and consumers of social media engage with argumentation. This information can then be applied towards creating tools and environments that discourage these types of abuse to facilitate more social argumentation.

1.2 Hypothesis and Research Questions

One key feature of social argumentation is the presence of an audience (Van Eemeren and Grootendorst, 2004; Jiménez-Aleixandre and Erduran, 2007). The audience's perception of the argument is something that is often overlooked in formal models of argument. Evidence suggests that perception of argument can be altered through multiple means such as cultural associations (Suzuki, 2011), pre-existing biases (Arceneaux, 2012) or peripheral information (Lee and Shin, 2014).

The ultimate aim of this research is to explore how a person's perception of argument on the social web, and their willingness to engage with it, differs based on the types of argumentation tactics used. To achieve this, it is first important to be able to correctly model and represent the arguments that occur socially. In this way, the key features of informal arguments can be identified and categorised. This can then be used to determine exactly which features of argumentation are considered most important by users, and those that they are most likely to engage, reply to, critique, and how these features shape users' overall interpretation of an argument.The work described in this thesis examines how formal models currently map arguments, and applies a combination of these models to an argument (or arguments) on the social web to determine which features are well captured, and those that are not.

This forms the basis of the hypothesis which is examined in the body of this thesis:

"A model of eristic argumentation on the social web should include both logical and rhetorical tactics, as the inclusion of rhetorical techniques affects the way in which users perceive and engage with the argument"

This is expanded through addressing four distinct research questions:

R1. *Is modelling eristic argumentation necessary for modelling argumentation on the social web?*

R2. *Are current frameworks and tools sufficient to model eristic argumentation on the social web?*

R3. *How should rhetorical techniques be included in a model of eristic argumentation on the social web?*

R4. *Do rhetorical techniques affect the way in which users perceive and engage with online argument?*

Question one is perhaps the most important question, as it determines the overall value of this work. It is answered in several different parts; firstly, by literature review, secondly, by an analysis of techniques commonly used in social argumentation, and thirdly by interviewing experts in fields that commonly use, model or support argumentation.

Question two focuses on determining whether it is currently possible to accurately describe argumentation occurring on the social web in terms of pre-existing models. Through a review of existing literature, the current state-of-the-art has been examined. Following this, a short exploratory work was carried out to evaluate the suitability of using one of these models to map examples of personal, social, and rhetorical argument from the social web.

Question three revolves around the most appropriate means of representing rhetorical tactics. Clearly, providing an exhaustive list of all possible examples of rhetorical tactics would not only be infeasible, but also unlikely to provide any value to modellers or analysts by simply overwhelming them with data. Therefore, to determine the most effective means of representing these tactics, modellers and analysts were consulted to determine the most effective method, with an emphasis on the purpose of use.

Question four focuses on the practical implications of this work; that is to say, whether the users of social media perceive arguments using different logical and rhetorical tactics in different ways, and whether this drives them to engage in different manners. This makes it important to define the terms perception and engagement. Perception can be thought of as the way in which users understand the tone, persuasiveness, entertainment value or information content of an argument (Sundar, 2000). Engagement, conversely, can be thought of as how likely users are to participate in the argument itself, and the different ways in which they do so. This not limited to simply replying to a post, for example: users of social media can engage in multiple ways, which include replying, sharing or voting (Markova and Petkovska-Mirčevska, 2013).

1.3 Report Structure

Background information on the topic area, both in argumentation and online behaviour, as well as the state of the research field at present, is discussed in Chapter 2. A preliminary investigation into the capabilities of current models of social argumentation, and an analysis of the results, is detailed in Chapter 3. In Chapter 4, these models are developed and adapted to encompass further social and rhetorical information, creating the Argumentation on the Social Web Ontology. This is used to examine the prevalence of a subset of rhetorical tactics in web-based argumentation and their correlation with machine readable features (such as post length, language, etc.). The model was developed further, with additional changes proposed, and review carried out in which experts in several relevant fields (argumentation modelling, linked-and-open data, the

social web, and philosophy) were asked to complete a pair of modelling exercises, once using the an existing argumentation, and once using the new additions. They were then asked a set of semi-structured questions about their experience. Chapter 5 details further data collection and annotation from sources on the social web, this time in the context of discussions surrounding online news. Again, this data is analysed at a structural level, in terms of both the social structure and annotated techniques. A narrative analysis is then carried out, examining three individual threads as case studies. In Chapter 6, the data gathered in Chapter 5 is used to form the basis of an experiment into the perception of argumentation: how logical and rhetorical techniques affect the perception, and reaction to, arguments on social media. Finally, Chapter 7 summarises the overall findings of this body of work, discusses the implications, and makes some suggestions to how this work can be expanded in future.

1.4 Contributions

The work discussed in this thesis has formed the basis of a number of papers:

- Blount, T., Millard, D. E., and Weal, M. J. (2014). Towards Modelling Dialectic and Eristic Argumentation on the Social Web. In *14th workshop on Computational Models of Natural Argument*

  This paper discusses the preliminary work carried out in Chapter 3, in which an existing model of argumentation is applied to a set of discussions on the social web, an its overall effectiveness evaluated.

- Blount, T., Millard, D. E., and Weal, M. J. (2015a). An Investigation into the Use of Logical and Rhetorical Tactics within Eristic Argumentation on the Social Web. In *ACM Conference on Hypertext and Social Media*

  This work forms the first part of Chapter 4, in which the Argumentation on the Social Web Ontology is developed, and trialled on a sample of argument data taken from the social web.

- Blount, T., Millard, D. E., and Weal, M. J. (2015b). On the Role of Avatars in Argumentation. In *Proceedings of the 2015 Workshop on Narrative & Hypertext*, pages 17–19. ACM

  This paper presents a position on one of the issues considered out of scope of the main body of work presented here: namely, do avatars—the visual representation of a person in a virtual world (Bailenson and Blascovich, 2004)—affect they way in which people argue, or the way in which they perceive arguments from others.

- Blount, T., Millard, D., and Weal, M. (2016). An Ontology for Argumentation on the Social Web: Rhetorical Extensions to the AIF. In *International Conference on Computational Models of Argument*

  This work concludes the work begun in (Blount et al., 2015a), developing the model further and presenting an expert review of the proposed changes. This forms the final part of Chapter 4.

Chapter 2

Background

2.1 Rhetoric and Argumentation

Rhetoric is often used in modern parlance as a pejorative, describing seemingly persuasive language that in reality lacks substance, or contains empty or insincere promises; formally, however, it refers to the art of persuasion, whether spoken or written. In particular, rhetoric focuses on the act tailoring one's argument to the situation at hand based on knowledge of events and, crucially, one's knowledge of one's audience (Corbett and Connors, 1999).

2.1.1 Modes of Persuasion

Aristotle, in his treatise on rhetoric, described three "persuasive modes" that can be employed in an attempt to sway an audience: through the words that are used (*logos*), through the character of the rhetor or their opponent (*ethos*), and through the emotions of the audience (*pathos*) (Kennedy, 1991). These modes may be applied individually, or in conjunction with one another. *Logos* describes an appeal to logic or reason. This is the method by which one might rationalise a position, often backing it up with evidence or statistics. It is important to note that, when enacting *logos*, it is not strictly necessary for the logic to be sound, or the evidence provided to be factual – it can be warped to fit a particular purpose, or even outright fabricated (however, this will usually also invoke another of the modes described below). The key element is that it appears to be reasonable and thus, appeals to an audience's sense of reason (Kennedy, 1991; Braet, 1992). *Ethos* is an appeal a person's character or sense of ethics and morals. This can be used in an attempt to strengthen the position of the rhetor's argument or to weaken their opponent's position. For example, if a rhetor can state that they are an expert in the field that they are debating then it is likely their audience will lend their argument more credence than if they were a novice. This specific case is known as an argument from authority, or *argumentum ab auctoritate* (Kennedy, 1991; Braet, 1992). Similarly, an argument can be made that attacks an opponents position indirectly, by attacking their credentials rather than refuting their claims

(*argumentum ad hominem*). Although such an argument is not logically sound (and constitutes a fallacy), it is still often used in practice and in certain circumstances is a viable (and often effective, if somewhat underhand) means of persuading an audience (Walton, 1987; Budzynska and Reed, 2012). Finally, *pathos* is an appeal to emotion, whereby an attempt is made to evoke a particular feeling in an audience in the hope that this will influence their opinion on a position. This can be done in both positive and negative terms. For example, flattering an audience, or promising them a boon, can shift them towards accepting a particular course of action. On the other hand, threatening them with the potentially undesirable consequences of their actions can cause them to reconsider even if these consequences are unlikely or, indeed, impossible. A classic example is the appeal to fear (*argumentum ad metum*) (Kennedy, 1991; Braet, 1992).

2.1.2 Dialectic and Eristic Argument

Persuasion is not the only reason for which argument occurs. The terms dialectic and eristic were coined in Ancient Greece to describe modes of argumentation with different goals, and were popularised in Plato's *Republic* (Plato, 80BC). A dialectic argument takes the form of two or more parties engaged in rational discourse with the aim of either discovering the particular truth behind a matter, or formulating a solution or resolution for a set of circumstances (Kerferd, 1981). For example, an academic presenting their findings and rationalising that they are indeed valid, given the rigorous methodology they have used and the weight of evidence this has provided is an example of a dialectic argument. Likewise, a peer reviewer that disagrees with the findings by pointing out a specific flaw in the experimental methodology and explaining how this should be resolved, is another example. The arguments tend to rely heavily on the mode of *logos*. In contrast, an eristic argument is an argument in which there is no clear resolution in the minds of the participants: they are not motivated by solving a problem, or convincing their opponent, but to be victorious (Kerferd, 1981). There may be different reasons for arguing in this vein, from quarrelling for its own sake as a form of catharsis (Schneider et al., 2014), to being seen to "win" the argument in the eyes of spectators (Jørgensen, 1998). As a result, these arguments chiefly favour the modes of *ethos* and *pathos*.

2.1.3 Modelling Argument

A common way to study how argumentation is used, as well as to aid the development of tools to encourage new ways to participate in argumentation, is to create a representation, or model, of the way in which an argument is carried out. Some of these methods rely on an underlying framework that specifies a formal theoretical basis of the rules required to create the models themselves.

These methods include incorporating notions of trust (Wigmore, 1913, p. 752), a focus on argument topic or chronology (Klein, 2010), and the ability to demonstrate support for, or refutation

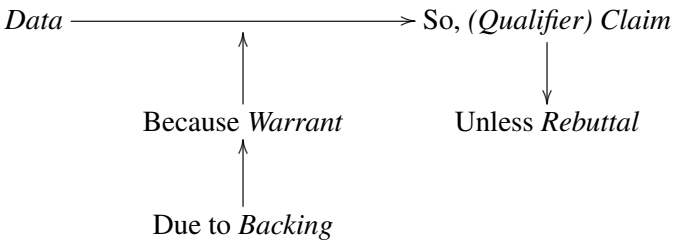


Figure 2.1: General form of Toulmin's diagram (Toulmin, 1958, p. 104)

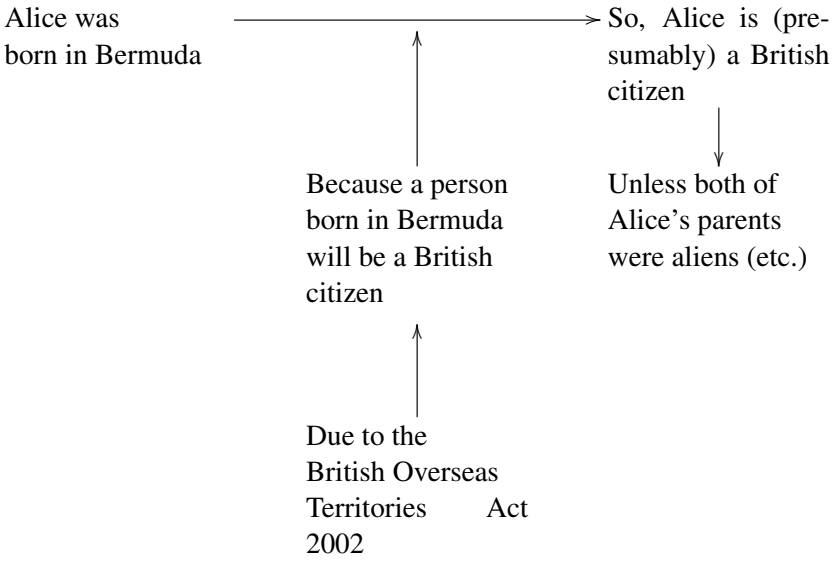


Figure 2.2: Example usage of Toulmin's diagram (Toulmin, 1958, p. 105), examining whether Alice is a British citizen

of, other points in the structure (Dung, 1995). Some examples of the means to model argumentation are discussed below, with respect to their technical structure, their influence in the field, and their practical applications.

2.1.3.1 Toulmin Model

Toulmin developed his model from the school of philosophy in the 1950s as a means of demonstrating an approach to practical (rather than theoretical) argumentation, by attempting to show the internal structure (and thus, consistency) of an argument (Toulmin, 1958). The general form of Toulmin's argument, shown in Figure 2.1, follows the structure of a *claim*, or conclusion, that is backed up with generally agreed upon facts (the *data*). The *claim* can be *qualified* ("definitely", "maybe", "probably", etc.) and any potential *rebuttals* accounted for. Then, key to the Toulmin model, the *claim* and *data* are connected using either an implicit or explicit *warrant*, or justification – this can then be supported by a particular *backing* (Verheij, 2005, p. 347-350). A specific example can be seen in Figure 2.2, which shows an argument reasoning that Alice is a British citizen.

Toulmin's model has been a particularly influential piece of work and has had an impact of decades of argumentation research in fields as far ranging as law, rhetoric and education (Newman and Marshall, 1992, p. 8-10; Schneider et al., 2013, p. 5, 12). However, there has been discussion as to the effectiveness of different aspects of the framework. In its favour, the means of explicitly stating the connecting warrant (and associated backing) can improve cross domain discourse. On the other hand, because models themselves are focused towards internal structure, there is no criteria for modelling overall structure (such as a group of arguments that refutes or support one another's claims). There is also no concept of resolving an argument (for example, on the grounds of logic or value); although this may have been by design, it negates the possibility of evaluating the strength of a given argument (Newman and Marshall, 1992, p. 349-350; Verheij, 2005, p. 5, 12).

Among other applications, the Toulmin model has been incorporated into the Argument Model Ontology[1], an OWL ontology to allow classification of academic arguments. This is used in conjunction with the Citation Typing Ontology (CiTO), an ontology for factually and rhetorically categorising citations (Peroni and Shotton, 2012, p. 8).

2.1.3.2 Argument Model Ontology and Citation Typing Ontology

The Argument Model Ontology (Peroni and Shotton, 2012) reifies Toulmin's (1958) model as an OWL ontology, using a set of familiar classes such as *Claim*, *Evidence*, and *Qualifier*, as well as properties such as *backs*, *involves*, *relates to*, etc. (Ruiz-Iniesta and Corcho, 2014).

This was developed in particular to be used in conjunction with the Citation Typing Ontology[2] (CiTO), an ontology for describing references between (and within) academic works. CiTO supports citations to other academic publications, as well as to external resources (such as data files that are hosted on the web), using *cites* (and reciprocal *isCitedBy*) relationships (Shotton, 2010).

However, citations in CiTO do not just form a link between two scholarly works: they may also model the rhetorical relationships between them. A positive citation can be used to show that a particular piece of work *confirms*, *credits*, or *supports* another, whereas a negative citation may be used when it *corrects*, *critiques*, or simply *disagreesWith* preceding work.

2.1.3.3 Scholarly Ontologies Project

The Scholarly Ontologies Project (ScholOnto), like CiTO, was developed as a means of supporting academic discourse through modelling scientific argument. However, whereas CiTO was designed to model the rhetorical links *between* documents and publications, ScholOnto was

[1] http://www.essepuntato.it/2011/02/argumentmodel
[2] http://purl.org/spar/cito

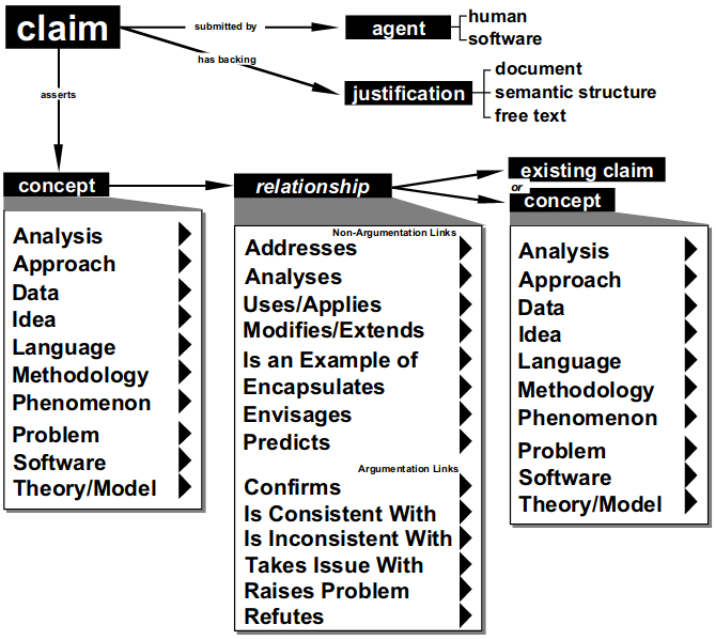


Figure 2.3: Anatomy of a Claim in ScholOnto (Buckingham Shum et al., 2000)

designed with the intent of modelling the arguments *within* them (Buckingham Shum et al., 2000).

Noting how the world wide web was fast becoming the primary means of publishing academic work, and desiring to keep the web true to its semantic origins (as envisioned by Berners-Lee), Buckingham Shum et al. (2000) developed ScholOnto as an ontology for representing scholarly claims, allowing academics to semantically mark-up their work in a form that could be automatically reasoned over or linked to other similarly marked-up papers.

In this way, authors can denote the contributions of their paper as a series of *Claims*, which represent one of a number of different concepts (such as an idea, a set of data, a methodology, a piece of software, etc.). *Claims* must have an owner (or *agent*), and some form of justification (such as an associated document). They may also share relationships with other concepts, such that one *Claim* may *Address*, *Analyse*, *Predict* or *Confirm* another. This structure is shown in shown in Figure 2.3.

An extension to ScholOnto, called ClaiMaker (Li et al., 2002), was also developed for the purpose of describing the rhetorical relations between claims in different papers (in a similar way to CiTO), by modelling them using relationships such as *is consistent with*, *agrees with*, or *refutes*.

2.1.3.4 Issue-Based Information Systems

Issue-Based Information System (IBIS) models are a particular type of dialectic process originally designed to aid in solving so-called "wicked problems" (Kunz and Rittel, 1970) – problems of social policy to which there is no clear definition, methodology or even end-goal (Rittel

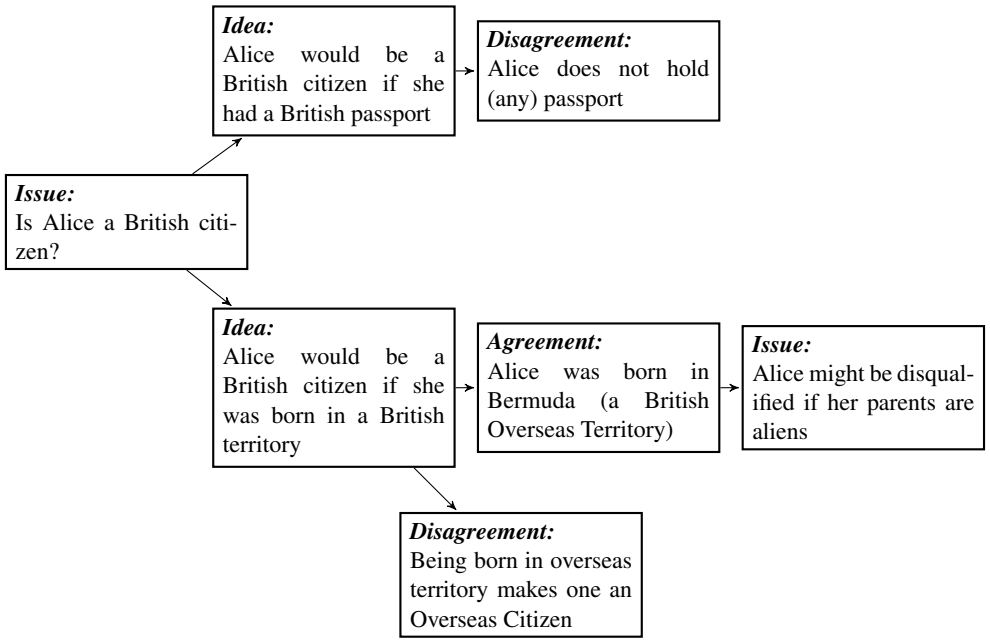


Figure 2.4: Example usage of an IBIS model, examining whether Alice is a British citizen

and Webber, 1973). IBIS models are represented as trees, made up of four different types of node. Firstly, *Issues* represent the problems that need to be solved, or questions that must be answered. Generally, there is one "root" *Issue* to be deliberated, but other sub-*Issues* can be created as necessary during the reasoning process. *Ideas* are proposed solutions or answers to these *Issues*, and each *Idea* can then be weighted positively or negatively using *Arguments For* and *Arguments Against*. IBIS models have seen wide usage in the field of design rationale and cognitive ergonomics where the assimilation of collective knowledge is required to solve problems (Conklin and Begeman, 1987; Aurisicchio and Bracewell, 2013). An example usage of an IBIS model is shown in Figure 2.4.

Because of its dialectic context, the application of IBIS models is ideal when two or more parties are trying to resolve a complex problem, especially if they have differing (or even opposing) stakes. As might be expected, there are many IBIS-like systems used in system-design and knowledge aggregation. Compendium[3] (Selvin, 1999; Selvin et al., 2001) is a hypertext "mind-mapping" tool developed to facilitate collaborative modelling, organisational memory and computer-supported argumentation. Through combining the IBIS principals of modelling exploratory dialogue with a more structured framework (including the use of predefined templates), Selvin et al. (2001) aim to reduce the "bottleneck" of knowledge capture and representation by supporting the process of knowledge elicitation during meetings in real-time, and tying existing organisation documents into the hypertext map.

[3] http://compendium.open.ac.uk/

Delibatorium[4] (Klein, 2010) is another tool that uses an IBIS approach to solving challenging problems, such as *"Is carbon offsetting a good idea?"*. In this, the IBIS approach is invoked to aid the collaboration of large amounts of people separated across space and time, by preserving a topic-centric (rather than time-centric) structure.

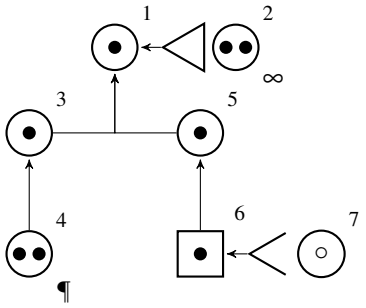
IBIS structures have also been included in an extension to the Semantically-Interlinked Online Communities (SIOC) ontology, devoted to representing argumentation occurring on parts of the social web (see Section 2.2.3).

2.1.3.5 Wigmore's Charting Method

"Wigmore's charting method", conceived of in 1913, is a means of recording argumentation originally devised for use in legal trials (Wigmore, 1913). Courtroom debate is an interesting area to consider modelling as it is an example of rhetorical argumentation in a formal setting. That is, there is a strict set of rules in place that governs the argumentation process e.g., participants may not be abusive, and each participant may only speak in turn, except to raise an objection, which may be sustained or overruled by the moderator (the presiding judge). There are often severe penalties that may be applied for beaching these rules (in the United Kingdom, the maximum penalty for "contempt of court" is a two-year jail sentence). Despite this, courtroom argument is a largely rhetorical process, in which prosecution and defence must "win over" the audience of jurors. Take, for example, the now famous line from the closing statement of the defence in the trail of OJ Simpson: *"If [the glove] doesn't fit, you must acquit"*. There is no reason for this statement to form a rhyming couplet, other than the impact it provides which was likely leveraged to make it stand out in the minds of the jurors.

Wigmore's method models the chain of interactions between competing arguments from each participant, and can be used to evaluate the validity of the overall conclusion that should be drawn (Wigmore, 1913, p. 751). The chart takes the form of a directed graph where each node represents a particular fact. The shape of each node relates to the nature of the assertion; squares represent testimony given under oath; a triangle represent an explanation of or support for the node it "points" to; an open angle refutes the argument it points to and all other assertions (such as claims, physical evidence or related legal statutes) are represented by circles. These can additionally be marked to denote arguments by the defence or prosecution, but are not discussed here for clarity (Chalamish et al., 2011, 2013). Symbols relate further information about the nature of these assertions: an infinity symbol (∞) states that a node denotes sensory evidence that may be (re)produced in court; a pilcrow (¶) denotes an assertion that can be taken as fact with no further evidence (such as a precedence case); a lack of a symbol shows that the claim is implied from further reasoning in the graph. In addition, Wigmorean analysis can incorporate the notions of *strong belief* (●●), *belief*(●) *doubt* (?) *disbelief* (○) and *strong disbelief* (○○) (Wigmore, 1913, p. 751-756; Goodwin and Fisher, 2000). An example structure is shown in Figure figure:wigmore.

[4] http://deliberatorium.mit.edu



1 Alice is a British citizen

2 Alice has a British passport

3 A person born in a British territory will be a British citizen

4 British Overseas Territories Act 2002

5 Alice was born in Bermuda

6 Alice's parents testify that she was born in Bermuda

7 Alice's parents' testimony could be biased in her favour

Figure 2.5: Example Wigmore graph, examining whether Alice is a British citizen

Little is known about precisely how often this type of analysis is used manually, although it is thought that it is carried out in courthouses around the world (Chalamish et al., 2011). However, efforts are being made to automate the process by parsing the natural language propositions made in court and transforming these into a Wigmore diagram to aid judges, barristers and juries in their deliberations (Chalamish et al., 2013).

2.1.3.6 Dung's Framework

Similar to Wigmore's method, Dung's framework (which uses the format of set theory) focuses on the aspect of arguments attacking, (implicitly) supporting and, ultimately, defeating one another (Dung, 1995). Dung defines an *Argument Framework* as a pair such that $AF = \langle AR, attacks \rangle$ where AR is a set of arguments $\{a_1, a_2, ..., a_n\}$ and $attacks$ is a binary relation such that $attacks \subseteq AR \times AR$. $attacks$ describes which arguments are "defeated" by one another: for example, if a_1 is the argument "Alice is not a British citizen" and a_2 is the argument "Alice has a British passport" then $(a_2, a_1) \in attacks$. The set of *conflict free* arguments is a maximal set of arguments that do not attack each other. An argument a_1 is *acceptable* with regard to a set of arguments S if there is no argument a_2 that attacks a_1 that is not itself attacked by an argument in S. A set of arguments is *admissible* if each argument is considered *acceptable* with respect to the set. The maximal *admissible* set is known as a *preferred extension* (Schneider et al., 2013).

There have been a number of extensions to this framework. Bench-Capon and Dunne (2002) have extended this framework to incorporate the idea of "value" or principle to arguments. When circumstances arise such that two possible resolutions to a dispute are equally valid, different audiences will have differing preferences based on the principles they feel are most important. For example, say that two solutions for combating crime are put forward: reading the general

public's private correspondence or an expensive social program of education and rehabilitation. If each has been proven to be equally effective, audiences that value minimisation of cost may favour the former whereas audiences that value individual privacy might choose the latter. Dunne (2016) also extended this particular framework to account for participants in a debate steadily increasing the volume of their voices, up until the point where they are no longer discussing in good faith and are simply trying to shout one-another down, ignoring all attempts at moderation.

2.1.3.7 The World Wide Argument Web

The World Wide Argument Web (WWAW or, simply, the Argument Web) seeks to provide a way of collating the vast amounts of unstructured debate occurring on the world wide web, and drawing them together into a coherent structure that can then be queried, augmented, visualised, and reasoned over (Rahwan et al., 2007a,b; Rahwan, 2008), as well as directly used to further ongoing arguments occurring on the web (Snaith et al., 2012).

As part of the development of the Argument Web, an ontology based on Walton's argumentation schemes (1996) was developed, and dubbed the Argument Interchange Format (AIF) (Chesñevar et al., 2006). The AIF is described in further detail in Section 2.1.3.8, and its extension (the AIF+) in Section 2.1.3.9.

The infrastructure of the Argument Web is supported by AIFdb (Lawrence et al., 2012), a database service that allows the storage and querying of Argument Web data in multiple formats, in an aim to reduce barriers between different branches of the argumentation ecosystem.

2.1.3.8 The Argument Interchange Format

The Argument Interchange Format (AIF) is a framework for representing argumentation as a directed graph (Chesñevar et al., 2006). Created as part of the Argument Web project (Rahwan et al., 2007a), which aims to link the concepts of natural language argumentation with abstract mathematical modelling (including capturing *"linguistically sophisticated manoeuvres"* (Bex et al., 2013)), the AIF is primarily a description, with specifications in a number of languages including RDF and SQL.

At its highest level, the AIF can be conceptually divided into an "upper" ontology and a "forms" ontology. The upper ontology consists of the building blocks of the argument structure, while the forms ontology applies context, for example, by differentiating between logical attacks based on faulty evidence, witness bias, or appeals to authority. A summary of the overall structure of the format is displayed in Figure 2.6. Data, claims and conclusions that make up the argument are modelled by Information nodes (I-nodes). There can be no direct relationship between I-nodes. Instead, there must be an intermediary Scheme node (S-nodes). These S-nodes are subdivided

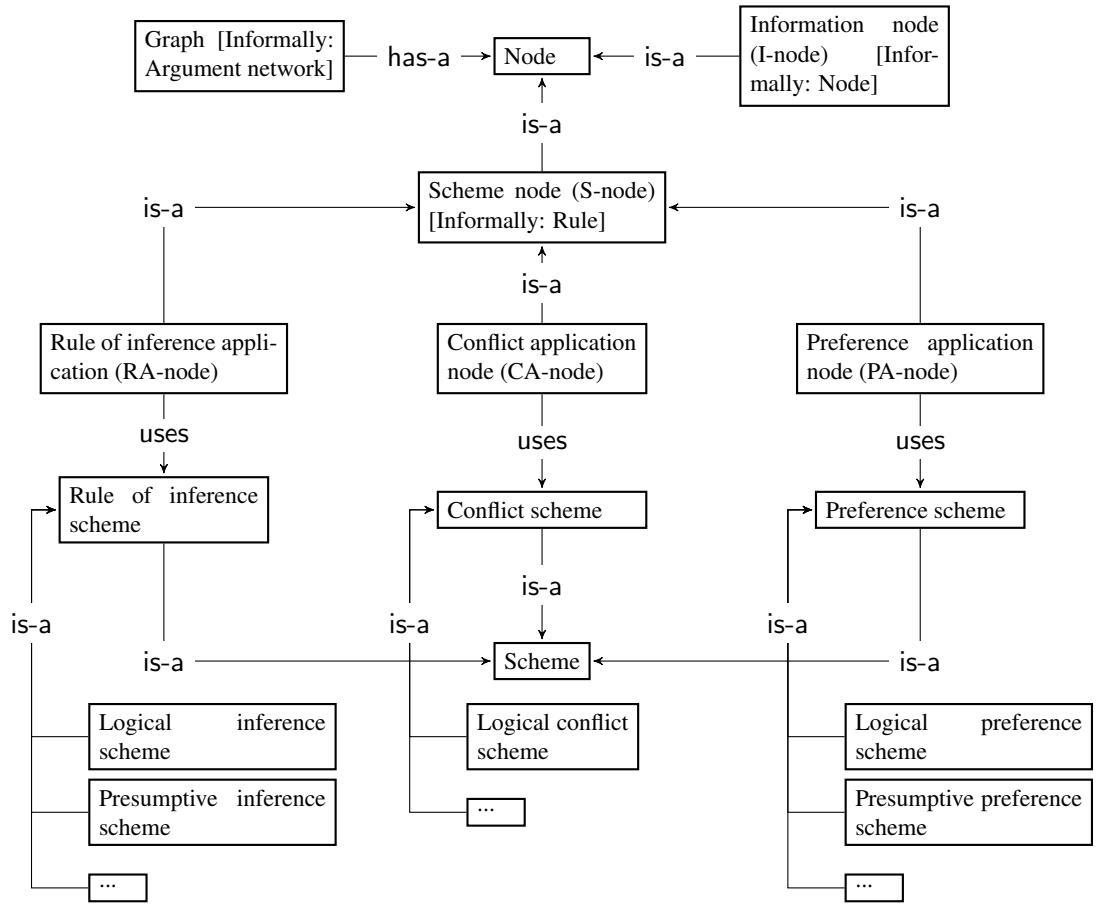


Figure 2.6: An overview of the AIF ontology, adapted from (Chesñevar et al., 2006)

into three applications: Rule of Inference Applications (RA-nodes), Conflict Applications (CA-nodes) and Preference Applications (PA-nodes). RA-nodes and CA-nodes simply denote an inference or conflict (logical or otherwise) between one or more pieces of information. PA-nodes, however, denote a preference of one piece of information over another. For example when discussing economics, while it may be difficult to logically prove the superiority of a regulated market over a free market, or vice-versa, the personal beliefs and preferences of proponent and opponent will feature heavily in their reasoning on such issues (Bench-Capon and Dunne, 2002). The implicit meaning of edges between these types of nodes is described in Table 2.1. The nodes themselves (and there appearance in diagrams throughout this thesis) are summarised in Table 2.2.

2.1.3.9 Speech Act Theory, Inference Anchoring Theory, and the AIF+

In their ongoing work on incorporating dialogue into the AIF, Reed et al. (2008) build on the work of O'Keefe (1992) to differentiate between two separate notions of argumentation: the first, which they term argument$_1$, is a logically constructed set of claims and evidence used to back these claims (or attack other claims), as in *"Alice put forward her argument"*. The second,

Table 2.1: Informal semantics applied to edges in the AIF, reproduced from (Chesñevar et al., 2006)

	to I-node	to RA-node	to PA-node	to CA-node
from I-node		I-node data used in applying an inference	I-node data used in applying a preference	I-node data in conflict with information in node supported by CA-node
from RA-node	inferring a conclusion in the form of a claim	inferring a conclusion in the form of an inference application	inferring a conclusion in the form of a preference application	inferring a conclusion in the form of a conflict definition application
from PA-node	applying a preference over data in I-node	applying a preference over inference application in RA-node	meta-preferences: applying a preference over preference application in supported PA-node	preference in supporting PA-node in conflict with another preference in PA-node
from CA-node	applying conflict definition to data in I-node	applying conflict definition to inference application in RA-node	applying conflict definition to preference application in PA-node	showing a conflict holds between a conflict definition and some other piece of information

Table 2.2: Summary of nodes in the AIF, as used in diagrams in this thesis

Name	**Description**	**Node**
I-node	**Information** nodes represent a (purported) piece of information, data, or a claim	
S-nodes (RA-, CA-, PA-nodes)	**Scheme** nodes denote a logical connection between I-nodes, respectively an **inference**, a **conflict**, or a **value preference**	

termed argument$_2$, refers to a dialogue—the exchange of ideas and opinions between two or more people, as in *"Alice and Bob were having an argument*.

This ties closely to the work on speech act theory (Austin, 1962, p. 108; Searle and Searle, 1969), which separates the notion of the locutionary act (the actual act of communicating, whether through words, writing or otherwise) with the illocutionary act (the intention behind them; to persuade, command, inspire or terrorise, for example) and the perlocutionary act (the actual effect of the words, whether intended or not).

Inference Anchoring Theory (IAT) (Budzynska and Reed, 2011a,b) provides a way to model the linkage between the inferred arguments$_1$ to the spoken (or otherwise communicated) argument$_2$. In IAT, the implied arguments$_1$ are anchored to the argument$_2$ through the addition of illocutionary force (although there is some debate as to whether these implicit speech-acts are actually explicit after all (Botting, 2015)). However, this illocutionary force can also apply to the transitions *between* arguments$_2$ in a dialogue through the nature of their content. For example, take the simple locution *"But why?"* in response to an assertion: the force behind this is a challenge, which is distinct from the force behind a response of *"Well, because of..."*, which provides a substantiation. Budzynska and Reed (2011a) describe these logical transitions between the participants as "dialogue glue", which allow for the modelling of (for example) "undercutting" of a particular position (negating it, without asserting the negation of it) based on (again, for example) their character in an *ad hominem* attack (Budzynska and Reed, 2012). Advances have been made in this field towards the automatic parsing of illocutionary structures within natural dialogue (Budzynska et al., 2016).

In parallel to the work on IAT, an additional set of nodes was introduced to the AIF in an extension dubbed the AIF+ (Reed et al., 2008). These new nodes are summarised in Table 2.3. The first, a subset of I-nodes dubbed Locutions (L-nodes), model locutionary acts (or utterances) in an argument$_2$. That is, they record precisely what was said. The second, a subset of S-nodes dubbed Transition Applications (TA-nodes), represent transitions between L-nodes (with associated forms such as a challenge or response). Thirdly, Illocutionary Applications (YA-nodes), also a subset of S-nodes, represent the "illocutionary force" and serve to link each argument$_1$ to the overall argument$_2$.

Figure 2.7 shows how this structure can be visualised. Consider the locution *"All men are mortal, and Socrates is a man. Therefore, Socrates is mortal."* The statement itself is modelled using the L-node on the rightmost side of the diagram. On the leftmost side is the core AIF structure, which show the premises formed as two I-nodes (*"Socrates is a man"* and *"All men are mortal"*), linked to the conclusive I-node (*"Socrates is mortal"*) by way of an RA-node. The L-node is connected to this argument$_1$ by way of the YA-node, shown in the middle.

Use of the AIF+, and the way in which it can be used to model argumentation specifically on the social web, is explored further in Chapter 3.

Table 2.3: Summary of nodes in the AIF+, as used in diagrams in this thesis

Name	Description	Node
L-node	**Locution** nodes represent the actual words that are spoken or written by participants	L
TA-node	**Transition** nodes represent links between locutions	TA
YA-node	**Illocutionary anchor** nodes tie the information and logical structure of an argument with the spoken or written locution	YA

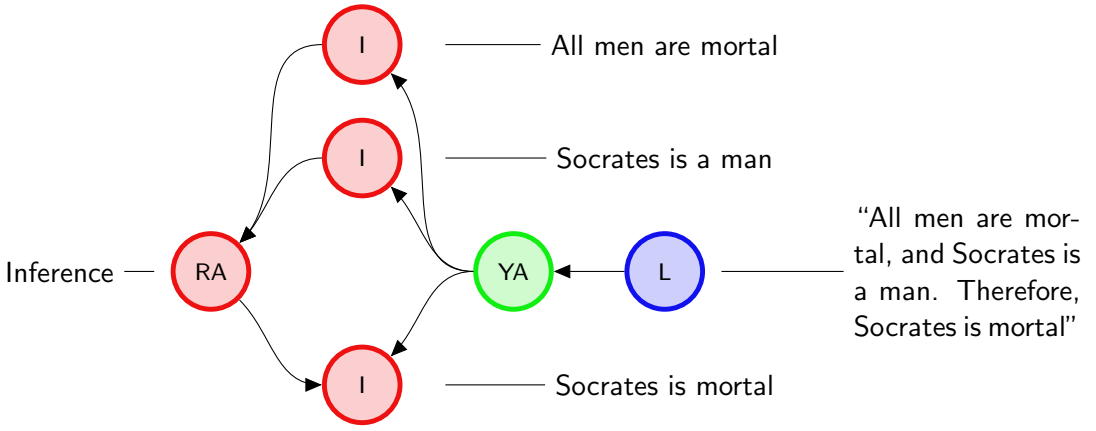


Figure 2.7: Visualisation of a simple AIF+ graph

2.2 Online Communication and Interaction

2.2.1 Social Media and the Social Web

The social web consists of the people, tools and communities that form over the world wide web, and is a way for individuals to share content, ideas and information. The social web presents a number of challenges for extracting and analysing arguments, particularly due to the lack of clear "indicators" of argument or structure. This problem is compounded by the type of language used; often highly informal, incorporating slang and irregular punctuation and grammar (Schneider et al., 2012). As the social web becomes more and more ubiquitous, the potential for using it to investigate how truly massive communities interact, communicate and argue increases dramatically.

Kaplan and Haenlein (2010) classify six distinct categories of social media: collaborative projects, blogs, content communities, social networking sites, virtual game worlds and virtual social worlds. Collaborative projects allow many different users to create, maintain and often discuss content. This category includes sites such as the online encyclopaedia *Wikipedia*[5], which

[5] https://en.wikipedia.org/

allow users to write and edit articles and *Urban Dictionary*[6], a user generated dictionary of slang and internet culture. Kaplan and Haenlein compare blogs (web-logs) to personal websites, in that they allow users to post information about the subject of their choice – these posts are often timestamped and presented reverse-chronologically. *Wordpress*[7] and *Blogger*[8] are two social media sites specialised for this purpose. "Micro"-blogging sites that pose limits on the amount of content that can be shared in a single post, such as *Twitter*[9], also fall into this category. Content communities revolve around the concept of publishing (and ultimately sharing) different forms of media. These include sites for publishing video (such as *Vimeo*[10]), images (such as *Flickr*[11]), audio (such as *SoundCloud*[12]) and many other different types of media. Social networking sites allow users to create a profile detailing information about themselves (such as home town, or music preferences) and then connect their profiles with the profiles of others on the site. Examples include *Facebook*[13] and *Google+*[14]. Virtual game worlds (such as *World of Warcraft*[15]) encompass online games in which a user controls a digital avatar to accomplish certain tasks (such as slaying a virtual dragon, or defeating another player's avatar). Similarly, virtual social worlds (such as *Second Life*[16]) encompass virtual spaces in which users have an avatar, but there is no specified aim or end-goal – the medium exists solely to facilitate social interaction. In this work, less focus is afforded to these latter two areas of the social web due to the the issue that as participants are controlling a virtual avatar, and may be playing a particular "role" rather than their real self, this can affect their behaviour and engagement in a discussion Hooi and Cho (2013). There is also the tendency for discussions to centre on the mechanics of the game world itself (Alagoz, 2013).

2.2.2 Anti-Social Behaviour

Anti-social behaviour is a growing problem on the social web, and often arises from debates or discussions that get out of hand (Suler and Phillips, 1998; Davis, 2002; Sood et al., 2012). This behaviour can arise from simple misunderstandings due to the difficulty in conveying tone through text, or as a deliberate act by individuals lashing out at other participants in a discussion. Incidents include flaming, in which a user simply hurls abusive language in an attempt to elicit an emotional response from their target, or to simply shut them out of the debate (Papacharissi, 2004; Konijn et al., 2008); spamming, in which a user floods the medium with content, often unrelated to the topic in hand, in the hope of drowning out other participants or as a means of advertising a commercial product (Krause et al., 2008); trolling, in which a user posts seemingly

[6] http://urbandictionary.com/
[7] http://wordpress.com/
[8] http://blogger.com
[9] http://twitter.com/
[10] http://vimeo.com/
[11] http://flickr.com/
[12] http://soundcloud.com/
[13] http://facebook.com
[14] http://plus.google.com/
[15] http://battle.net/wow/
[16] http://secondlife.com

innocuous but deliberately fallacious argument to provoke other members of the group into becoming outraged (although there is debate as to whether this term refers to the bridge-dwelling monster of myths, or the fishing term for dangling a baited line behind a boat) (Herring et al., 2002); and much more serious incidents of directed harassment, threats and stalking (Spitzberg and Hoobler, 2002; Willard, 2007; Jane, 2014).

As a result, there is a concerted research effort into the best way to tackle these issues before they cause serious harm to individuals, or the field as a whole. Suler and Phillips (1998) discuss a wide variety of approaches (specifically in regard to the virtual social world *The Palace*[17], but these could be applied to other online spaces as well). The simplest solution is to moderate users' interactions and dispense warnings, "mutes" (where a user may observe, but not contribute) or, in extreme cases, bans as and when the situation warrants. While effective for dealing with small or close-knit communities, this approach does not scale when considering the social web.

A different approach is to allow the community a degree of self-moderation. Reputation systems, for example, allow users within a community to assign "votes" to a particular account, or post, to show its trustworthiness. This allows new users to make judgements on whether to take a comment seriously, for example, or to purchase something from a particular seller in an online auction (Resnick et al., 2000; Anderson et al., 2012). However, this can also lead to a feedback loop in which communities become self-reinforcing; if users always vote for posts of similar sentiment (or against those that disagree), then gradually these sentiments will become dominant. Over time only users who hold these views will contribute to the site (further reinforcing the disparity) and the community as a whole will stagnate or worse, become distrustful or outright hostile to new members or "outsiders".

In another example of direct self-moderation, the popular online game *League of Legends*[18] implements a "tribunal" system in which players that are reported for poor behaviour in matches (such as verbally abusing team-mates) are judged by their peers. These peers can examine evidence such as chat logs and game scores, then decided whether to "pardon" or "punish" the offending player (Hodson, 2013; Kou and Nardi, 2013; Blackburn and Kwak, 2014).

A more covert attempt to manipulate users' behaviour can be found in certain implementations of human-computer interaction design. HCI can be leveraged to "trick" users into performing (or not performing) an action desirable to the designer. These so-called "malicious interfaces" (Conti and Sobiesk, 2010) are often used to trick users into spending time or money that they otherwise would not (for example, advertising banners that suddenly cover page content). In 2008, YouTube temporarily added an "Audio Preview" button to its comment system that would read aloud what the user intended to post[19]. This was placed in the previous place of the "post" button (which had been moved further to the right), such that a user was likely to unintentionally preview their comment before posting it (Munroe, 2008).

[17] http://thepalace.com
[18] http://leagueoflegends.com
[19] Likely inspired by this *xkcd* comicstrip: https://xkcd.com/481/

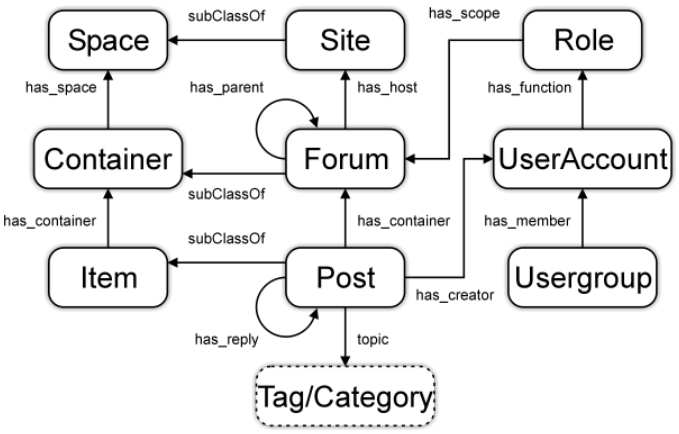


Figure 2.8: An overview of the core SIOC ontology[20]

2.2.3 Semantically-Interlinked Online Communities

The Semantically-Interlinked Online Communities project (SIOC) aims to enable the cross-platform, cross-service representation of data from the social web (Breslin et al., 2006), and is widely adopted when modelling user-generated content (Schmachtenberg et al., 2014). SIOC allows for semantic representations of *Sites*, which hold *Forums*, which contain *Posts* (which *has_reply*), authored by the owner of a *UserAcount*. This structure is shown in Figure 2.8. SIOC is often used in conjunction with the Friend of a Friend (FOAF) ontology, to show how individuals map to their online accounts.

An extension to SIOC for the purposes of capturing and representing argumentation was developed by Lange et al. (2008) based on the principles behind Issue-Based Information Systems (as discussed in Section 2.1.3.4). Under this scheme, SIOC *Posts* can be annotated as being a *Statement*, using a subclass of *has_reply* when one statement *refers_to* another. Similar to the original model, these *Statements* can be subclassed as an *Issue*, an *Idea*, an *Elaboration*, or a *Position* (when discussing a personal preference). This structure is shown in Figure 2.9.

Two use-cases are outlined by Lange et al. (2008) in the initial proposal. The first is to account for "forum and blog discussions", in which an open question is raised (such as "where should we go to lunch on Sunday"), participants pose suggestions, and subsequently raise objections or approval, along with their reasoning. The second use case covers "wiki discussions and bug tracking", in which a known issue is raised (such as a page that violates a specific policy of the wiki, or a bug in a particular version of software) and can be elaborated on, before concrete courses of action are proposed to resolve it.

However, both of these use cases are framed in terms of *discussion* rather than *argumentation*; that is, they share the assumptions that participants are contributing in good faith, and that there will be, ultimately, a resolution to the discourse, a decision will be reached, and the issue at hand will be solved satisfactorily (more or less) for all involved. Therefore, while this approach

[20] http://sioc-project.org/ontology

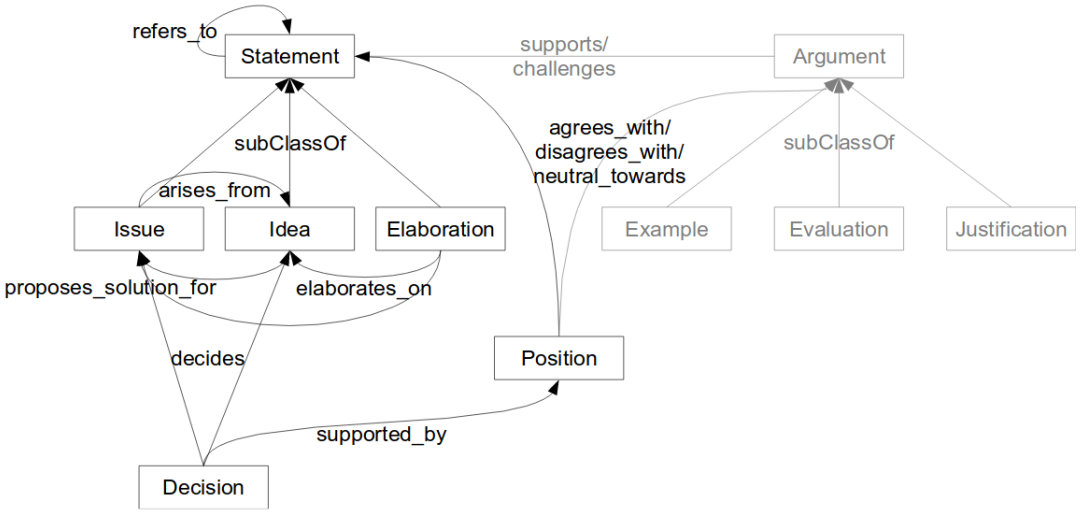


Figure 2.9: Lange et al.'s (2008) argumentation extensions to SIOC

highly useful when dealing with social media discourse centred around deliberation (and to a lesser extent, criticism or inquiry), it is not necessary a suitable approach when considering modelling eristic arguments.

2.3 Social Aspects of Argumentation

2.3.1 Modelling Argumentation on the Social Web

Many theoretical models of argumentation are based on the assumption of a dialectic argument, as their purpose is to aid the participants with the process of understanding the information discussed, or to reason over the model and draw conclusions regarding the outcome. However, in social media there is a clear proliferation of eristic argumentation (Sood et al., 2012).

It is also valuable to consider the effect that anti-social behaviour can have on a discussion. It is important to separate the notions of civility and politeness when considering argumentation: while often used interchangeably, conflating them in this way can lead to ignoring the *"democratic merit of robust and heated discussion"* (Papacharissi, 2004). This highlights the case that the use of arguments that may be considered "rude" within a discussion does not necessarily diminish their importance when considering their impact on the debate. Indeed, anti-social behaviour has been shown to further polarise the viewpoints of participants in a debate (Anderson et al., 2014).

This makes the role of audience an important feature to consider: when an individual responds to a post on the social web their post is often seen not just by the author of the post they reply to, but by many other users as well. In fact, many posts may be directed at this wider audience to seek approval, voice dissent, or provoke other emotions (Berland and Forte, 2010). Equally,

the tone of the discussion can be changed by the participants perception of the way the audience feels about particular issues (Anderson et al., 2014). Consider the analogy of a political hustings: neither candidate believe they can change the mind of their opponent, but instead are debating with a view to sway their audience; as a result, their positions will often take on a rhetorical nature, taking into account what they know of the audience to better persuade them to their point of view.

Schneider et al. (2014) note though, that currently it is difficult to model the value of eristic arguments as participants are free to *"sling propositions that they would not commit to under other circumstances"* as a means of catharsis, recreation or entertainment. In addition, because argumentation is a social and subjective process, it is also important to recognise the fact that individuals may perceive the same argument in many different ways due to cultural beliefs (Suzuki, 2011), pre-existing cognitive biases (Arceneaux, 2012), as well as features surrounding the content of the argument such as avatars (Lee and Shin, 2014).

2.3.2 Social Argumentation Tools and Platforms

There are many websites forums and discussion groups set up specifically to encourage their users to engage in discussion and argumentation. One example of these is Reasonwell[21], a collaborative argument map that allows users to argue for recreation, setting forth arguments with one another in a formalised manner, such as explicitly contributing claims and providing evidence for them. Similar to this is Cohere, a web-based tool for knowledge aggregation, idea linking, and argument visualisation (Shum et al., 2008). Users can enter prepositions or Ideas, and then link them with labelled connections, specifying whether the connection is positive, neutral or negative, in the same fashion as when using the IBIS model described in Section 2.1.3.4.

Torroni et al. (2010) also highlight the need for including social features in models of argument on the social web, going so far as to outline proposals for an Authoritative Social Web Platform (AeSoP) in an attempt to combat trolling and the spreading of falsehoods across social media. Torroni et al. hope that by combining the implementation of an argumentation framework that integrates argumentation features (such as claims and attacks, similar to Reasonwell) into existing social media platforms, in conjunction with direct community feedback (such as "Likes" or up-/downvotes), they will be able to highlight (using colour, spacial positioning, ordering, etc.) posts that are deemed "social acceptable" arguments, to create an environment conducive to reasonable debate while maintaining the social (and informal) atmosphere that users expect of the social web.

ArguBlogging[22] (Snaith et al., 2012; Bex et al., 2014) provides an approach to blogging inspired by the semantic web, and provides a way of connect their arguments across the web to the Argument Web itself. It achieves this by providing a "bookmarklet" interface, in which users

[21] http://www.reasonwell.com/
[22] http://argublogging.com

can select text from the webpage they are browsing (such as a news site, a blog, or encyclopedia page) and respond directly to the passage they have selected with whether they agree or disagree, and the reasons they have for doing so (as well as a title for the generated (micro-)blog), before posting it to one of two linked blogging platforms with an embedded "Argue" button, that allows a later reader to respond in the same fashion. The subsequent chain of debate can then be viewed as part of the Argument Web through, for example, AIFdb.

Similar to this is Online Visualisation of Argument[23] (OVA+) (Janier et al., 2014), a tool built on the foundation of the AIF, allowing a user to create argument maps from the social web (and the web in general) *in situ*, by highlighting text from the page and marking them up directly as elements of the AIF(+) (as described in Sections 2.1.3.8 and 2.1.3.9) such as I-nodes, RA-nodes or CA-nodes, and the relationships between them, and contributing it to the Argument Web.

2.4 Summary

To begin to answer the research questions $R1$ and $R2$, this chapter examined the existing literature in the fields of argumentation, and online social interaction and behaviour.

In answer to $R1$ (*Is modelling eristic argumentation a valuable direction of work?*), after exploration of literature examining argumentation as a social process (and the consequences eristic argumentation can have on individual discussions, platforms, and augmentation as a whole), it is reasonable to state that examples of this type of argumentation are indeed important to consider; even if rhetorical argumentation cannot be reasoned over in the same fashion as formal or dialectic arguments, the role it has in shaping a discussion and the impact of its presence on an observing audience should not be ignored.

To begin to answer $R2$ (*Are current frameworks and tools sufficient to model eristic argumentation on the social web?*), a number of frameworks used to model argument were examined. The majority of these are deliberately prescriptive, rather than descriptive, such that (by design) they do not take into account eristic argument or rhetorical tactics. This means that although they are well suited to modelling argument that can be formally reasoned over, and can even aid in the construction of a well-formed dialectic argument, they may struggle to model common forms of argumentation on the social web.

In the following chapter, to strengthen the answers to $R1$ and $R2$, an existing framework for modelling argumentation is applied to a real example of social dialogue, to determine the prevalence of (and feasibility of modelling) eristic argumentation on the social web.

[23] http://ova.computing.dundee.ac.uk/

Chapter 3

Modelling with the Argument Interchange Format

Having begun to answer in Chapter 2 the research questions $R1$ and $R2$ (addressing the value of modelling eristic argumentation, and the ability of existing models to adequately represent it) through a review of existing literature in the field, this chapter seeks to establish through a practical application of existing modelling frameworks both the prevalence of eristic argumentation on the social web (and hence the additional contributions that could be captured through a socio-rhetorical-focused approach), and the ability to accurately model social argumentation and the nuances between dialectic and eristic argumentation, with existing tools.

To this end, a preliminary investigation was conducted, in which an existing framework for modelling argumentation (the Argument Interchange Format (AIF), as described in Sections 2.1.3.8 and 2.1.3.9) was used to model a real example of argumentation taken from the social web. The AIF was determined to be the closest fit-for-purpose ontology for modelling argumentation on the social web, due to the goals of capturing practical, language-based argumentation, with the additional benefit of being readily extensible. This exercise aimed to show whether the AIF can be combined with other ontologies in a way that makes it easier to model rhetorical argument, and to determine the key strengths and weaknesses of this combination in relation to modelling social argumentation.

3.1 Methodology

3.1.1 Combining the AIF and SIOC

To explicitly capture the social component of argumentation on the social web, while also modelling the formalised argument structure, the key elements of the AIF+ (summarised in Table 3.1) were combined with those of the SIOC ontology in this approach (the SIOC ontology being

Table 3.1: Summary of nodes used in conjunction between AIF and SIOC, as used in diagrams in this thesis

Name	Description	Node	
I-node	**Information** nodes represent a (purported) piece of information, data, or claim		AIF
S-nodes (RA-, CA-, PA-nodes)	**Scheme** nodes denote a logical connection between I-nodes, respectively an **inference**, a **conflict**, or a **value preference**		AIF
L-node	Social media Posts are treated as individual **Locution** nodes, and represent the actual words written by participants		AIF+
TA-node	**Transition** nodes represent links between **Locutions**		AIF+
YA-node	**Illocutionary anchor** nodes tie the information and logical structure of an argument with the spoken or written **Locution**		AIF+
U-node	**UserAccount** nodes denote the social media account the user uses to contribute		SIOC

widely adopted when modelling user-generated content (Schmachtenberg et al., 2014)). This is achieved by linking the concept of a SIOC Post directly with that of an AIF Locution (L-node), treating each social web thread as a separate dialogue, or argument$_2$, and considering each post as an atomic unit within the dialogue (containing zero or more individual arguments$_1$).

In the majority of cases, a single locution will translate to a single self-contained argument$_1$. However, it is also possible for a single post to contain a number of arguments$_1$, each with a number of premises and a single conclusion. In this situation a single L-node will link to multiple YA-nodes, as shown in Figure 3.1. In rare cases (often caused by constraints imposed on the length of a post by the service, such as the 140 (since raised to 280) character limit on Twitter), a user will spread the premises of a single argument across multiple posts to construct their argument$_1$. Figure 3.2 shows how, in such a situation, multiple L-nodes will link to a single YA-node. If two users post identical statements, they still contribute two distinct locutions. However, they will both be linked to the same I-node(s), and therefore the same argument$_1$. In this situation, multiple YA-nodes may point to the same I-node, such as in Figure 3.3.

3.1.2 Data Collection

A single topic of argumentation was chosen to be examined for three case studies, each representing a different social media system. To ensure the stimulation of debate, the selected topic

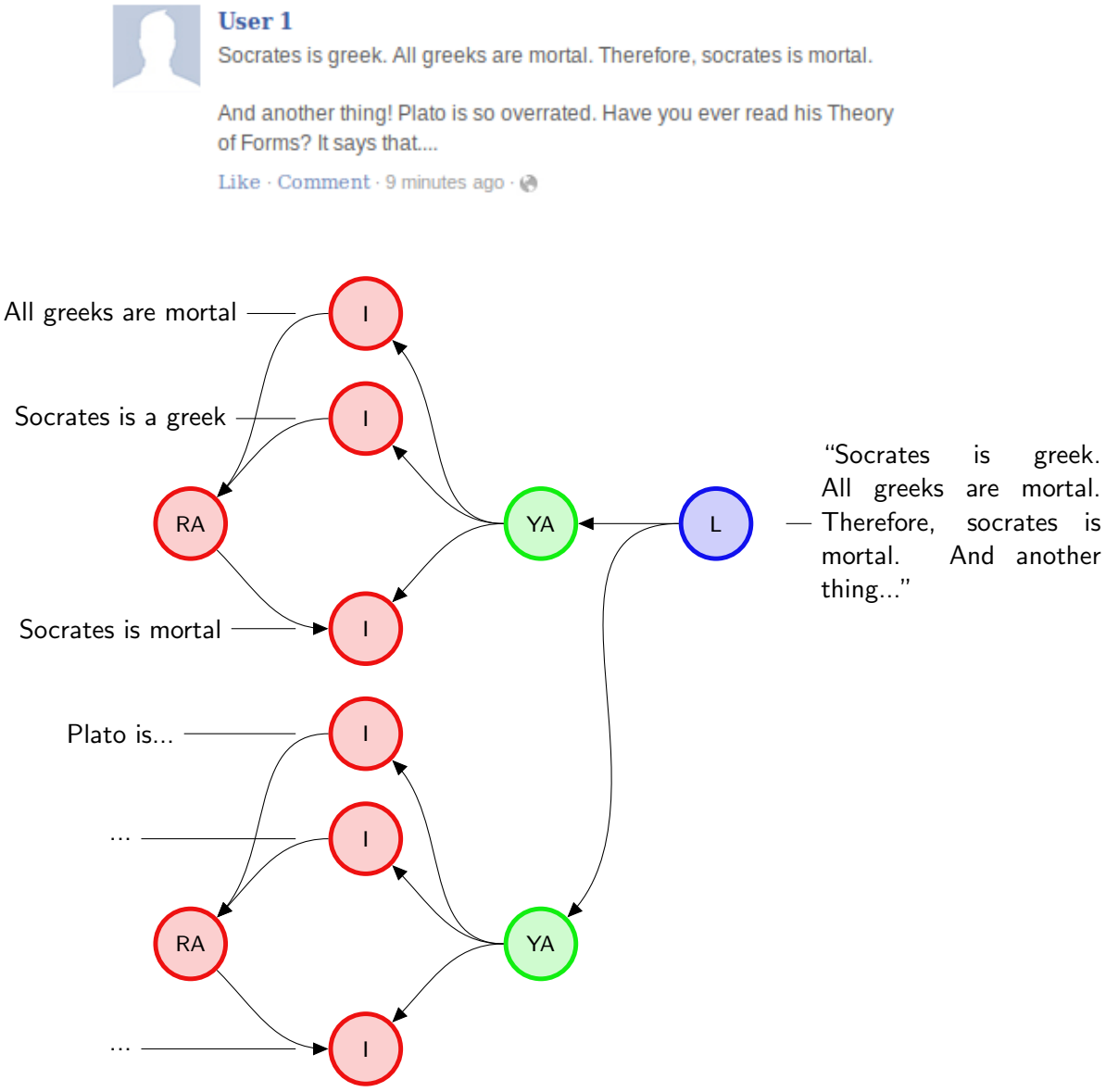


Figure 3.1: Visualisation of one post making two distinct arguments[1]

needed to be controversial, have a large number of respondents and have been active for a long enough period of time to generate a rich and complete content. The October 2013 United States government shutdown caused by Congress's failure to agree on a budget, and the following condemnation this received from the presidency, was a suitable match for these requirements.

This topic was then tracked across three of the social media categories identified by Kaplan and Haenlein (2010): Twitter, a microblogging service that allows users to publish messages of up to one-hundred and forty characters; Facebook, a social network, that allows users to create a network of "friends" and share text or images; and YouTube, a content creation site where users can create and upload videos, or playlists of videos.

The source of the posts themselves again needed to be both publicly available and have a large

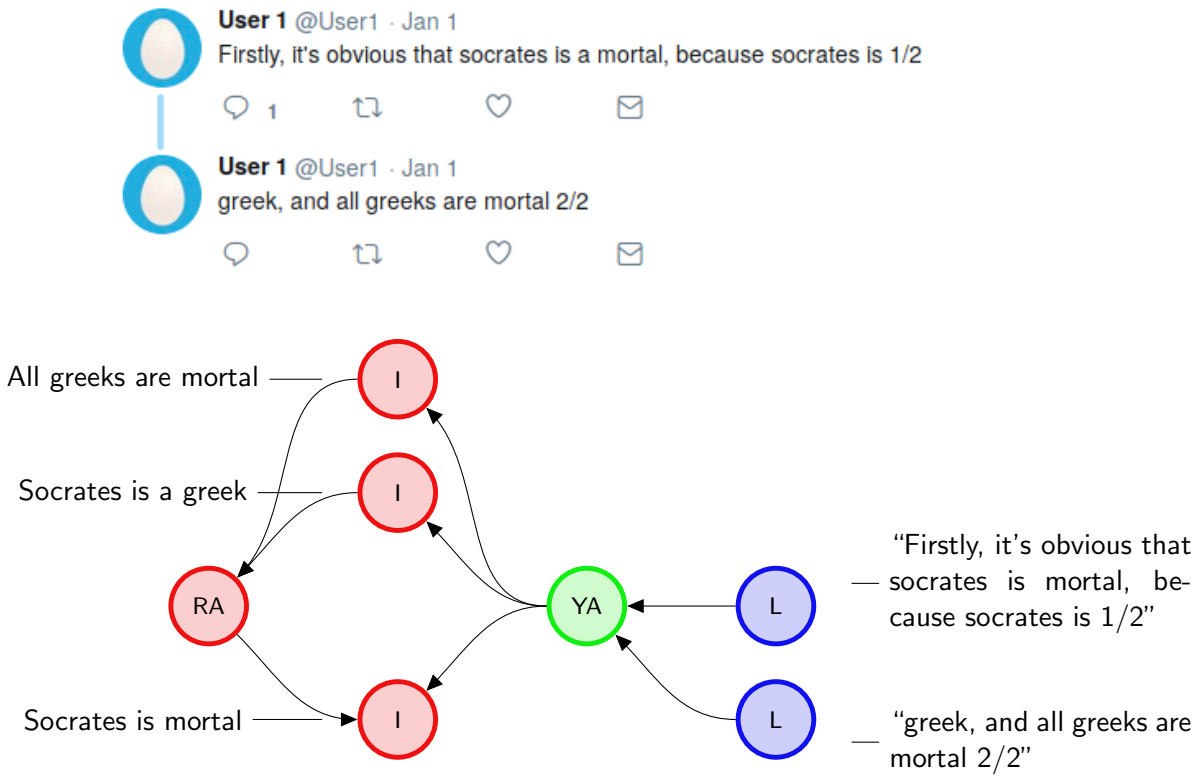


Figure 3.2: Visualisation of two posts, used to construct a single argument$_1$

number of followers to ensure a maximally stimulated debate. As an authoritative public figure at the heart of the crisis, content from or relating to Barack Obama's social media profiles was chosen, and three posts that were broadly similar in content were selected for study. The first post, initially posted on 8th October 2013 from the White House's YouTube channel[1], is a 14m 40s video recording of Obama delivering a statement to press from the West Wing of the White House, condemning the shutdown. The post taken from Obama's official Twitter account[2] (which is managed by a third party, Organizing for Action), dated 15th October 2013, reads: *"This is unacceptable. Tell Tea Party Republicans to stop holding our economy hostage: http://OFA.BO/qNmA3Y"*. The included hyperlink leads to an Organising for Action page, which encourages users to to voice their displeasure at the shutdown by allowing them to automatically generate and send tweets. The post taken from Obama's official Facebook account[3] (also managed by Organizing for Action), also dated 15th October 2013, reads: *"Tea Party Republicans in the House of Representatives forced a government shutdown, and now they're threatening an economic shutdown. This has gone on for too long. Tell them to #EndThisNow: http://OFA.BO/ACC7qB"*.

The discussions surrounding these posts were acquired by collecting comments replying to each initial post, and those replying to subsequent posts in the discussion (taking into account only

[1] https://www.youtube.com/watch?v=7LwoudGfug0
[2] https://twitter.com/BarackObama/status/390288744235823104
[3] https://www.facebook.com/photo.php?fbid=10151874920756749

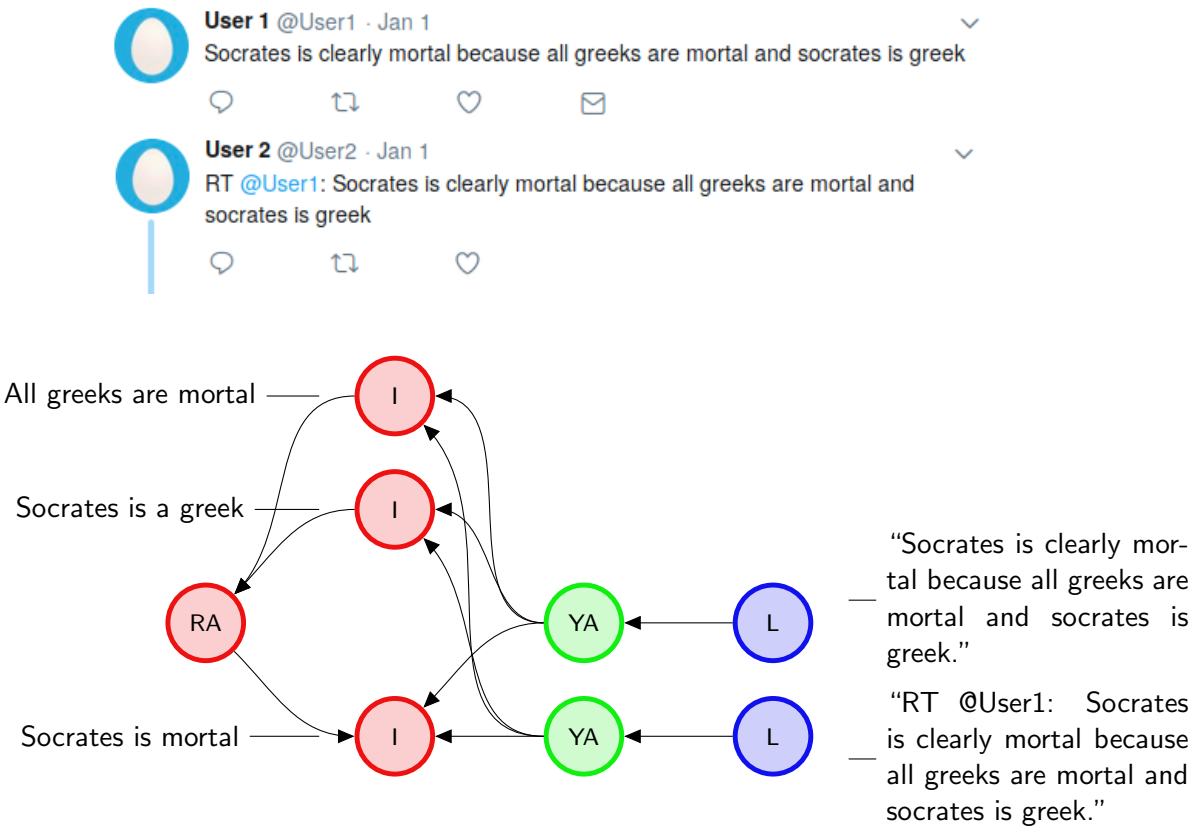


Figure 3.3: Visualisation of two posts, repeating the same argument[1]

Table 3.2: Metrics of total dataset collected from YouTube, Twitter and Facebook

Metric	**YouTube**	**Twitter**	**Facebook**
Total number of posts	2719	137	9494
Total number of users	1255	33	6224
Degree (longest chain of replies)	2	22	0
Average posts per user	2.17	4.15	1.53
Average words per post	26.74	15.91	40.12
Average characters per post	150.13	97.63	241.14
Time between first and last posts	101d 16h 19m 12s	0d 13h 40m 48s	90d 19h 55m 12s
Average time between posts	53m 52s	3m 02s	13m 47s

direct replies, rather than mentions within the text of the post), with the use of the public Twitter, Facebook and Youtube APIs respectively. This data was translated to an RDF triple-store using SIOC to record the data specific to the social media platform, such as which User created which Post and which Thread stores which Posts. This was used in conjunction with the DCTerms ontology, which held supplementary data such as timestamps.

3.1.3 Data Sampling and Annotation

Because of the volume of the data produced over the course of the tracked event and the time-intensive nature of manually annotating the data, it was necessary to sample the data to a more manageable size before annotation could take place. To prevent information being lost when the dataset was scaled down, it was important to ensure that the sampled graph maintained properties (such as diameter and average path length) similar to those of the raw data (which in the context of social media threads refer to the distribution of replies). To maintain these characteristics, "forest fire" sampling (Leskovec et al., 2005; Leskovec and Faloutsos, 2006) was used to create a sub-graph that preserved the overall structure of the parent by attempting to maintain a similar degree distribution and diameter.

The algorithm for forest fire sampling is as follows:

1. Choose a "forward burning probability" p (in this instance a value of 0.7 was chosen based on the recommendation by Leskovec and Faloutsos (2006) for scaling down a larger graph)
2. Choose a random starting node
3. Add this node to the sample graph. Select x nodes at random from all nodes linked to the chosen node, where x is a random number geometrically distributed with mean $\frac{p}{1-p}$. If the selected node has fewer than x linked nodes, select all available nodes, and return to step 2.
4. With each selected node, recursively repeat step 3 until the desired sample size has been reached.

The full networks described in Section 3.1.2 and Table 3.2 were scaled down to thirty replies (not including the original posts), using this method. The effect of the sampling on the metrics recorded is shown in Table 3.3. Broadly, the gathered metrics are similar, with the exception of time between posts, suggesting the forest fire algorithm samples posts from the whole scope of the threads. While this does introduce the possibility that some context is lost, this is an acceptable trade-off to reduce the likelihood of sampling small clusters of argument within the thread that differ from the overall structure.

The sampled data was then manually annotated by a single researcher (the author) with the formal argument$_1$ information. Specifically, from each L-node, both explicit and implicit I-nodes were extracted and related together using the most appropriate S-nodes.

For example, Obama's original Twitter post (an L-node) states: *"This is unacceptable. Tell Tea Party Republicans to stop holding our economy hostage: http://t.co/y8fPF8s3bG"*. From this the following I-nodes can be extracted: *"The Tea Party Republicans are holding the economy hostage"*, *"Holding the economy hostage is an unacceptable tactic"* and *"The Tea*

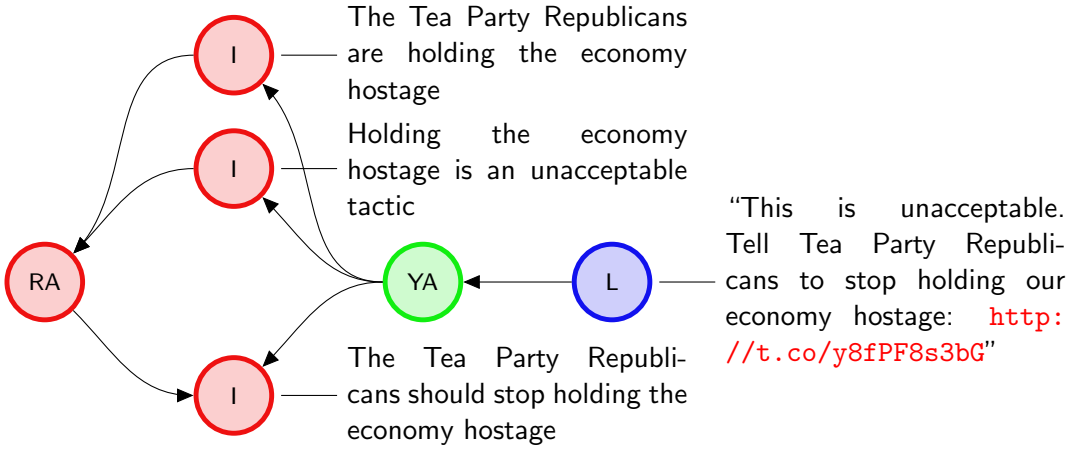


Figure 3.4: Mapping of Obama's social media posts to the AIF

Table 3.3: Metrics of discussions sampled from YouTube, Twitter and Facebook

Metric	**YouTube**	**Twitter**	**Facebook**	**All**
Total number of posts	30	30	30	90
Total number of users	23	12	30	65
Degree (longest chain of replies)	0	16	0	16
Average posts per user	1.30	2.50	1.00	1.38
Average words per post	26.77	16.33	42.10	33.18
Average characters per post	147.90	101.20	259.67	201.70
Time between first and last posts	4d 0h 54m 56s	0d 5h 13m 33s	3d 12h 13m 18s	n/a
Average time between posts	3h 20m 31s	0h 10m 49s	2h 54m 15s	0h 17m 10s

Party Republicans should stop holding the economy hostage". From this, it is easy to see that the single locution contains two premises and a conclusion (which therefore need to be joined using an RA-node). This argument$_1$ can then be mapped to the specific locution by means of a YA-node, as shown in Figure 3.4.

3.2 Results and Analysis

An overview of the raw data collected from each platform is shown in Table 3.2 and the sampled data in Table 3.3. In total, the discussion generated by the Twitter post has slightly over one-hundred and thirty replies—in contrast, the YouTube comments total nearly three thousand posts, and the Facebook discussion has well over nine-thousand. Each platform sees the vast majority of posts contributed soon after the initial post. However, each has a "long tail" of responses that gradually decrease in frequency as time goes on. The discussion on Twitter seems particularly ephemeral, with participants only contributing for a short time before moving onto other topics; while the Facebook and YouTube posts appear more "permanent", with users finding and contributing to them months later.

Table 3.4: Aspects of raw data from social media APIs capable of being modelled using the AIF or SIOC ontologies

Features present in social media APIs	Represented in: AIF	SIOC
Locution (explicit content)	✓	✓
Illocution (premises/conclusions)	✓	
Argumentation structure (attacks/support)	✓	
Author	✓	✓
Avatar		✓
Replies	✓	✓
Creation Date	✓	✓
Reputation (e.g. "Likes")		
Location		
User "Type" (i.e. individual/business/etc.)		
Sentiment (implicit content)		

In addition, when collecting this data it became apparent there was information that had no appropriate representation in either ontology, such as reputation systems (for example, the "Likes" used by Facebook), the sentiment of the post (for example, sarcasm, humour, abuse) or information about the type of user making the remark (whether they are an individual, a celebrity, a corporation, etc.); these omissions are shown in Table 3.4. These features could have substantial bearing on the perception of the argument$_2$. Consider the example of reputation systems: a retort stating *"You're an idiot"* may be perceived very differently by the audience if it has no up-votes, one up-vote or one hundred thousand up-votes. Alternatively, consider a user making the argument$_1$ that *"I really love using this product"*: whether the statement is made by an individual, or the company selling the product would likely influence the validity and value of the statement.

Table 3.5 shows the statistics collected after annotating the data with premises and conclusions, represented as AIF nodes. Given this data it can be seen that Twitter is the only sample that contains Transition-nodes; that is, replies to other posts within the thread. While this may appear to suggest that the platform is used more for debate than the others, it is possible this is down to deficiencies in the APIs of the other platforms, which may not accurately expose replies. It can also be observed that the debates on Twitter and Facebook have a much higher information content than that of YouTube.

On the surface, the sample of posts taken from Twitter and Facebook appear to have similar information content. However, upon manual inspection, it can be seen that this average is actually heavily skewed by one particular Facebook post that is thirteen paragraphs long and contains a total of twenty six information nodes. The argument in question is reproduced on a number of different websites, and is likely reused in full as a boilerplate "cut and paste" rebuttal by many users when engaging in an argument on that topic.

To highlight the overall information disparity take, for example, the tweet *"@BarackObama Stop expanding government, spying on Americans and driving up the deficit."*. This is an

Table 3.5: Summary of AIF nodes found in annotated discussions collected from YouTube, Twitter and Facebook

Metric	**YouTube**	**Twitter**	**Facebook**	**Total**
U-nodes	23	12	30	65
L-nodes	30	30	30	90
TA-nodes	0	20	0	20
YA-nodes	31	30	41	102
I-nodes	88	116	110	314
S-nodes	13	30	26	69
L- to I-node ratio	15:44	8:29	3:11	45:157

enthymeme—the literally derived I-node acts as a conclusion, while the premises (that Obama is expanding government, spying on Americans and driving up the deficit, and that to do each of these things is detrimental) are left implicit. In turn, contrast with the posts *"first"*, *"wow obama"* and *"lolollll i love this"* which contain very little information, either explicit or implicit. In addition, not all posts with a large amount of literal content have a comparatively large amount of information. For example, posts such as *"Give DIRETIDE Give DIRETIDE Give DIRETIDE..."* (repeated upwards of fifty times in a single post) show a desire to derail the discussion by flooding it with completely irrelevant information ("Diretide" refers to a cancelled seasonal event in the popular online game *Defence of the Ancients 2*; the cancellation sparking uproar from the fanbase, which led to a number of social media platforms being flooded with this message). While seemingly nonsensical, even given the context, and completely unrelated to the original post on on-going debate, this post still makes a contribution to the discussion and colours the perception of those that see it (even if they simply respond by rising to the bait, and berating the user, rather than engaging with the rest of the discussion).

In addition, there are other posts that have deeper contextual meaning that would first appear. Consider, for example, "RedScareBot"[4]: this is an automated Twitter account, using the avatar of Joseph McCarthy (an American politician infamous for, at the height of the Cold War, making claims that their were numerous Soviet agents embedded in the US government), which replies to any tweet that includes phrases such as "communism" or "commie" with quips such as *"Commie Chameleon"*, *"Oh noes, Socialism"* or *"Rise of the USSA"*. While this may seem nonsensical or a *non sequitur* without context, *with* context it can be viewed by the audience as a derisive or satirical retort to a knee-jerk insult, despite being posted by a machine.

An overview of the resulting annotated structures are visualised in Figure 3.5, which shows a side-by-side comparison of the three different samples. A selection of examples are examined in more detail. In these diagrams, U-nodes are omitted for the sake of clarity. Some posts were relatively straightforward to model within the existing AIF. The post from Twitter, shown in Figure 3.6 for example, follows a standard premise-inference-conclusion structure. However, other posts are more problematic. Take for example the Facebook post shown in Figure 3.7; while the internal inferences can be adequately modelled, what of the emotive language used

[4] https://twitter.com/RedScareBot

(calling the Tea Party "creeps")? While this might be used as an attack on any contributions made by a spokesperson for the Tea Party, what if there are no contributions to attack? This highlights that it may be useful to model attacks on *entities* as well as their contributions. For an even more extreme example, consider Figure 3.8: what information could be retrieved from this tweet? Vulgarity aside, there is no logical contribution being made here, despite the impact that this response would undoubtedly have on the tone of the discussion. Alternatively, Figure 3.9 shows a much more pleasant response, on the surface, while still containing no real information. However, without additional context, it is difficult to say exactly the intention behind it: is it a genuine "smile", or is it sarcastic mocking? Figure 3.10 shows a Youtube post making a more obviously humorous post; again, however, although some information can be captured and represented through the AIF alone, the persuasive element inherent in using a joke (if any) is lost.

There are of course limitations on the conclusions that can be drawn from a relatively small dataset when working with proverbial "big data". As such, these findings cannot be used to justify broad claims that state that *all* arguments on a particular example of social media are structured in this way. These examples instead serve to demonstrate the important fact that different types of structures *can* evolve, and provide some examples of the argumentative and rhetorical tactics people use when arguing over social media and how the conjunction of the AIF and SIOC projects (as well as any extensions made to these) can be used in attempts to map them.

3.3 Summary

To trial modelling social argumentation with the AIF and SIOC, a small case study was performed, examining the practical implications of annotating a particular topic of argument on three social web platforms: Facebook, Twitter, and YouTube. This preliminary work shows that, currently, it is insufficient to use the AIF (and its extension) to fully model eristic argumentation, even when certain social aspects are modelled through other ontologies such as SIOC, due to the frequency with which "illogical" or information-less contributions are made to the discussion (that nevertheless will have an impact on the way the discussion as a whole is interpreted).

$R1$ and $R2$ can now be answered definitively. The prevalence of eristic arguments within the samples examined here strengthens the suggestion that there is value to modelling eristic argument simply due to the frequency with which they are deployed, and (without further augmentation) current frameworks appear to be, by design, unsuitable for modelling social argumentation (particularly as it appears on the social platforms examined in this chapter).

To address this, in the following chapter a number of augmentations to the AIF are proposed, in order to aid the more complete modelling or argumentation on the social web.

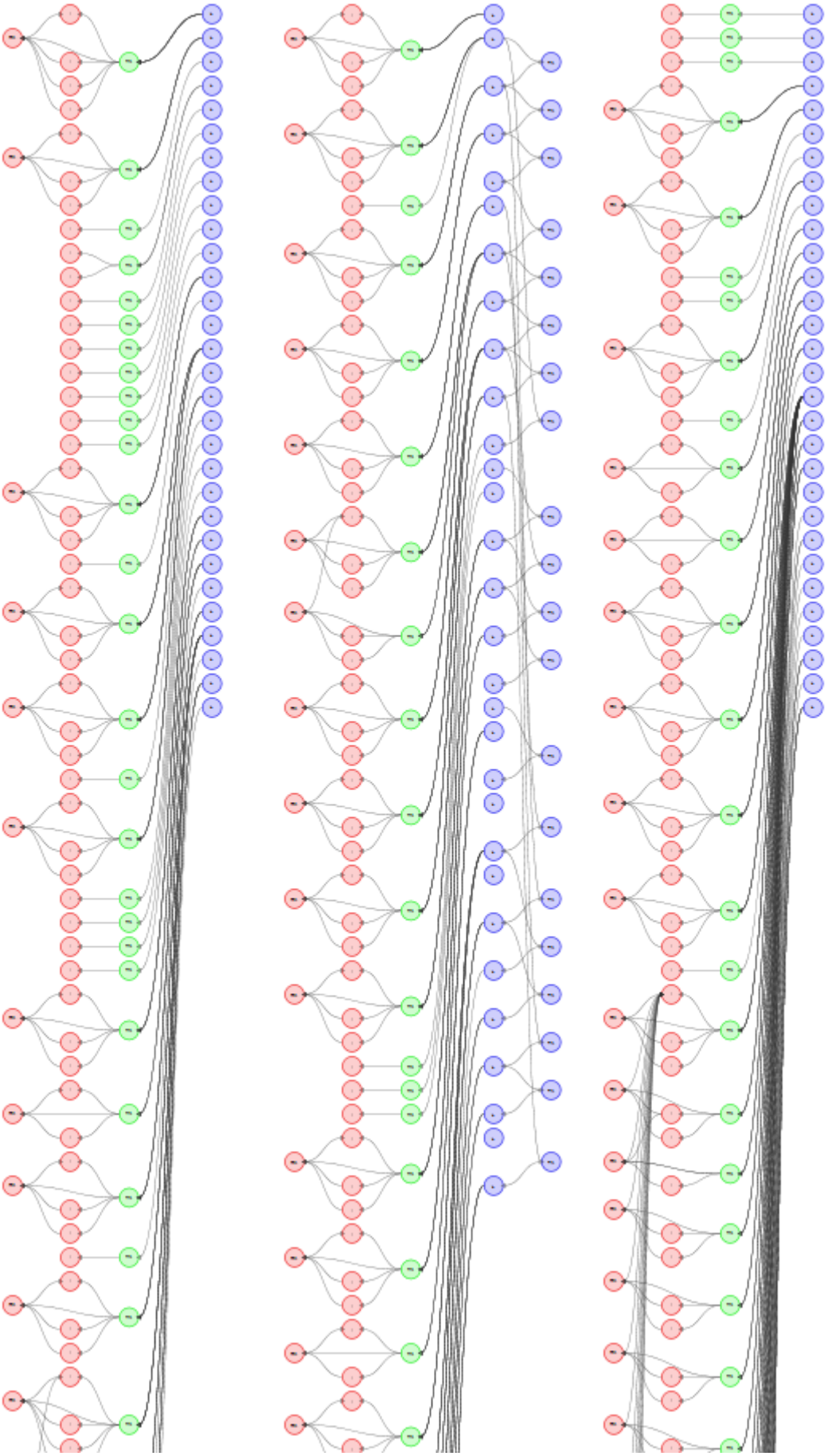


Figure 3.5: A side-by-side comparison of the emergent structures of discussions taken from YouTube (left), Twitter (centre) and Facebook (right)

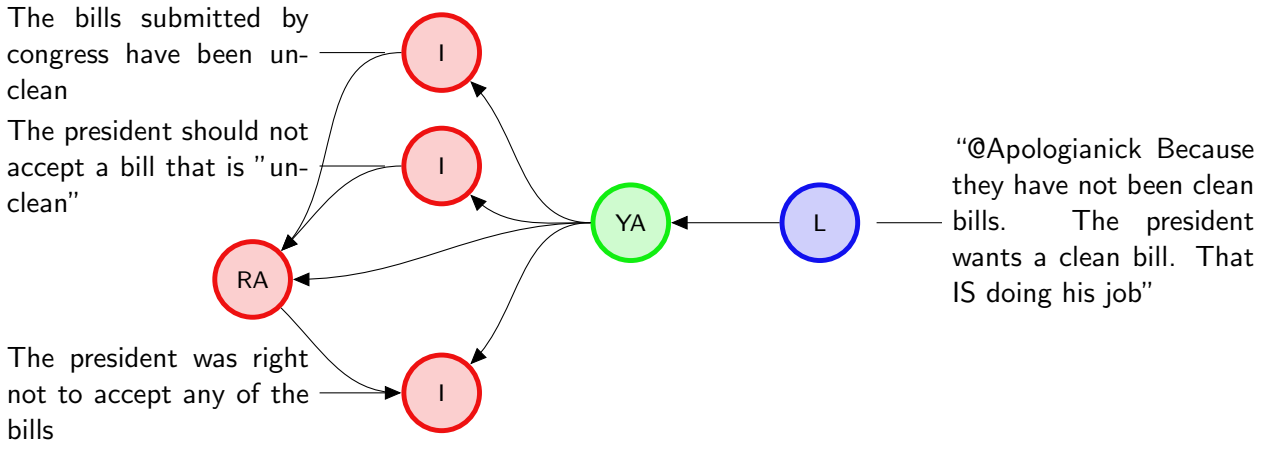


Figure 3.6: An example Twitter post, showing a standard premise-inference-conclusion structure

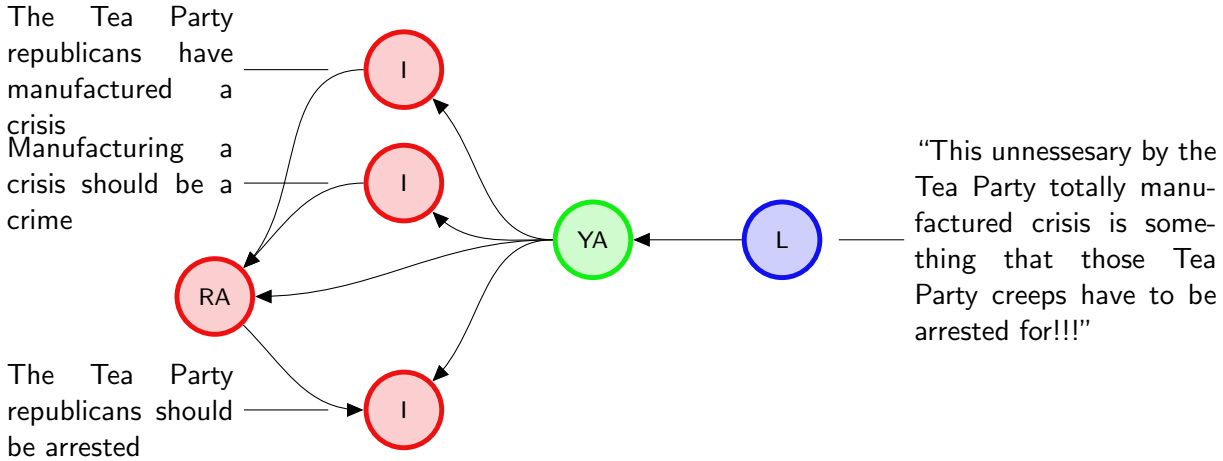


Figure 3.7: An example Facebook post, showing a moderately hostile response

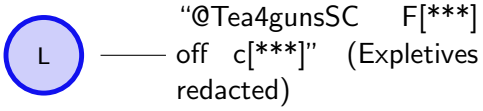


Figure 3.8: An example Twitter post, with no informational content, showing an extremely hostile response

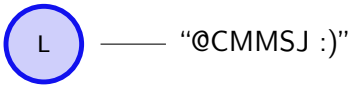


Figure 3.9: An example Twitter post, also with no informational content, showing a more pleasant response

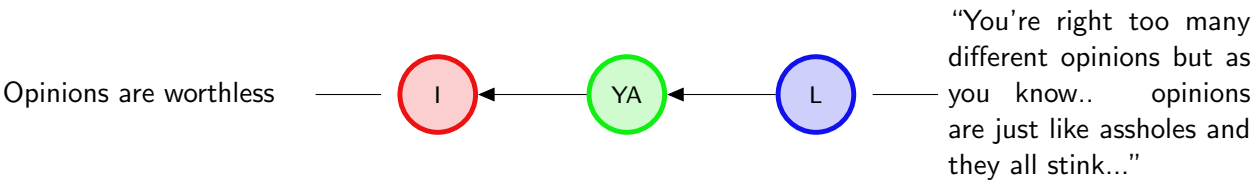


Figure 3.10: An example Youtube post, with minimal informational content, showing a humorous response

Chapter 4

The Argumentation on the Social Web Ontology

In the preliminary investigation carried out in Chapter 3, the capability of existing frameworks and their use in capturing and modelling argumentation and social communities was examined and evaluated (Blount et al., 2014) in order to answer research questions $R1$ and $R2$. It became apparent that the AIF, while a powerful tool for modelling dialectic argument, lacked the ability to capture the eristic aspects of social argumentation. This chapter aims to address research question $R3$, and answer how social (and eristic) argumentation could (and perhaps how it should) be modelled, by providing a proposal for an extension to the AIF that encompasses the ability to model socio-rhetorical elements of argument, and then refining these proposals through applying these extensions to a larger sample of the case study described in Section 3.1.2, as well as through an expert review process.

While some informal logical fallacies, such as the *ad hominem* attack, can be suitably modelled within the AIF (Budzynska and Reed, 2012), the impact on the debate of contributions that are no more than "simple" abuse is, ironically, more difficult to capture. There is reason to suggest that while, for example, abuse may not be valuable to the argument itself, it is still valuable to model such outbursts due to their overall effect on the discussion. A heckler in a debate, for example, may not have any well-reasoned argument to hand and resort simply to throwing vulgarities, but by merely disrupting the proceedings they are voicing their dissent at the positions offered. This is reason enough not to discard the contribution; however, it can also act to catalyse further argumentation on the subject between the main participants. Likewise, a participant in a debate may, instead of putting forth their own argument or attacking their opponent's, make some sort of joke to endear themselves to the audience and to lift their own standing, with the hope that this also raises the acceptability of their arguments. While the AIF can model the locution, the rhetorical force behind it—the effect of this type of emotive tactic on the debate as a whole, and the social pressures it entails—goes uncaptured.

In addition, there are other socio-rhetorical tactics that are often employed on social media. These include spamming (posting large volumes of a repetitive nature) to drown out other posters, deliberate deviation from the topic at hand, bringing up non-sequiturs in an attempt to derail the argument, and "meta-argumentation"—criticising the way in which an opponent argues, but not the argument itself (e.g. if a user claims another is breaking the rules of the forum, or of not arguing in good faith). There are also the non-textual features of social media to consider, including multimedia posts (such as image memes), or the the feature of posts other than their content, such as the number of "Likes", "Favourites", or "Retweets" a post has, which can demonstrate popular (or audience) support for that opinion, and subsequently influence the perception of that position.

4.1 Extending the AIF

As shown in Chapter 3, and presented at the Workshop on Computational Models of Natural Argument (Blount et al., 2014), the existing formal structure of the AIF (even combined with SIOC) is insufficient to fully model eristic debate. Therefore, a number of additions on how these ontologies could be adapted to model the socio-rhetorical aspects of argumentation proposed. While logical tactics must conform to principles of validity, rhetorical tactics are not bound by such constraints. They are not required to be internally or externally consistent, and instead are concerned with the use of social pressures, and how the audience can be used to influence the course of the debate. The principal focus here is the inclusion of rhetorical support and attack. Although these tactics are only one aspect of rhetorical argument, they feature heavily in eristic dialogue (particularly rhetorical attacks), showcase both the positive and negative aspects of rhetorical argument, and are important due to the impact they can have within discussions on the social web and the culture surrounding it (Blount et al., 2015a). Rhetorical support and attack (also referred to as *ethotic* support and attack) has also recently been incorporated into argumentation mining techniques that use a natural-language approach (Duthie et al., 2016).

Rhetorical support is often relatively benign, and can be used to show solidarity with other members of the dialogue, to incorporate oneself into a social group, or to encourage participants to put forth positions that are seen as favourable. Consider the extracts *"bro fist bump"*, a short declaration of support for another user (encouraging them that their views are accepted by the watching audience), and *"I commend you for admitting that debt & deficits are important...If only more* [people] *felt the way you do"*, which disagrees with the overall stance presented by their opponent, but commends them for conceding some common ground, in attempt to encourage further dialectic argument.

Conversely, rhetorical attacks are often (although not always) extremely hostile. They differ from logical attacks by emotively attacking the person behind the argument, rather than the argument itself, using social pressure to influence the debate. This is not to be confused with an *ad hominem* argument, which attacks a person's argument by calling their character or credentials

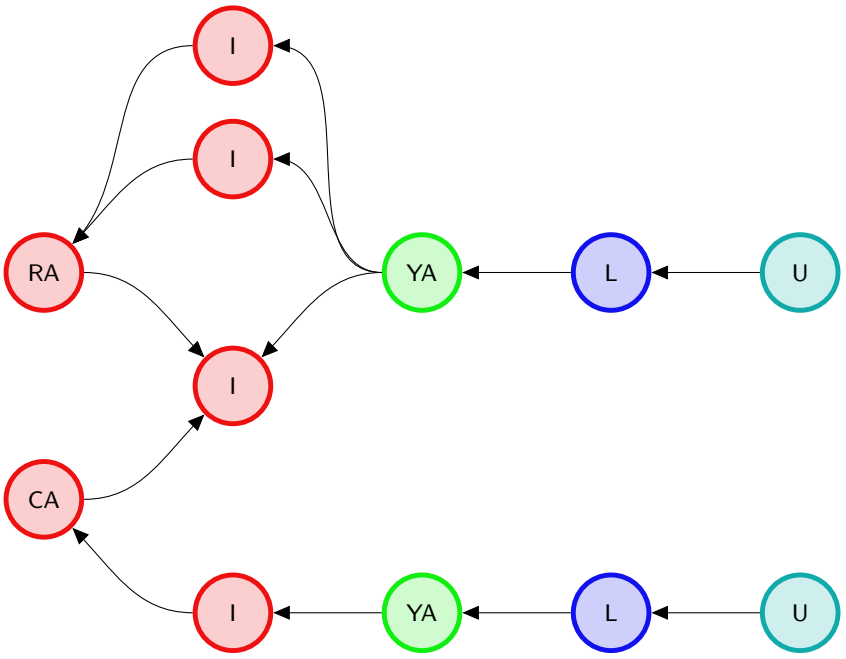


Figure 4.1: Proposal for representing abusive attacks as solely within the argument structure

into question—these are "logical", even though they are fallacious. Because rhetorical attacks rely on social pressure to make their impact, they often contain extremely vulgar language. The purpose of these statements can be interpreted in a number of ways, from intentionally silencing dissenting voices with threats and the deliberate construction of a hostile environment, to showing the audience how impassioned and emotive the rhetor is on the subject, or even to simply use aggression as a means of catharsis, relieving anger or stress.

Figure 4.1 shows the simplest approach, similar to the current way the AIF models the use of *ad hominem* attacks, by linking the attack to the opponent's argument with a CA-node. However, this is insufficient for the majority of abusive attacks; while *ad hominem* tactics attack an opponent's argument by claiming they are not qualified, or otherwise unfit, to make such an argument, abuse often does not attack their position at all, but seeks to undermine them emotionally in front of their peers.

This mapping can be modelled by linking the content of the locution to the targeted user's account as shown in Figure 4.2. However, a UserAccount can be involved in any number of topics, and be attacked for any number of reasons. Furthermore, a person can choose to present themselves as a dramatically different person (having different credentials, skills, opinions or even a different race, religion or gender) when they conduct themselves on the web, as opposed to in person. They may also choose to represent themselves differently between individual threads and discussions. As such, there must be some way to model a participant's (presented) character (or *ethos*) within the ASWO.

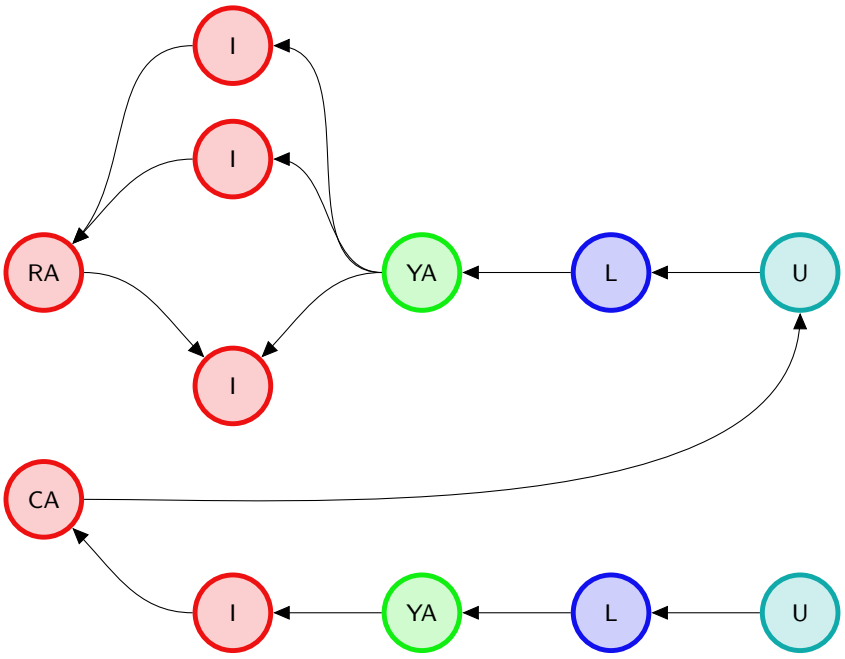


Figure 4.2: Proposal for representing abusive attacks as connected with the social aspect of the argument, attacking the author directly

To this end, the notion of a "Persona" node (P-node) is introduced. This represents the character that a participant assumes during a particular discussion. For example, a person may argue in a different manner in a debate about their preference of music than they do about their expertise of software engineering, political beliefs, or preference of restaurant. Likewise, they may argue on a given topic in a totally different fashion when in different company, tailoring their use of tactics to their audience, whether friends, colleagues, or complete strangers, or a mixture of each. To accommodate this, UserAccounts are permitted to be associated with many Personas where necessary, as shown Figure 4.3. Introducing the idea of a Persona allows each UserAccount to present a different view of themselves that, importantly, can be supported or attacked directly by argument nodes in the discussion, when engaging in multiple discussions or topics. However, when and where this particular structure is applied is likely subjective and based on the annotators' interpretation of what constitutes a different topic or a distinct Persona, and is therefore left to their discretion. The inverse—associating one Persona to multiple UserAccounts—is also possible, and could be used to represent the situation of a participant in an argument attempting to artificially solidify their position by creating multiple accounts (known as *sock-puppets*), and an example of a possible way of modelling this is shown in Figure 4.4. Again, the circumstances under which this can be applied (if, for example, they have enough evidence to conclusively decide that multiple accounts are managed by a single individual) are very much at the annotators discretion.

Using this model of a person's *ethos*, participants in a discussion can now attack (or support) one another's Personas directly, as shown in Figure 4.5. Here, the view of Budzynska and Reed (2012) (that in the AIF an *ad hominem* attack amounts to the "undercutting" of a particular

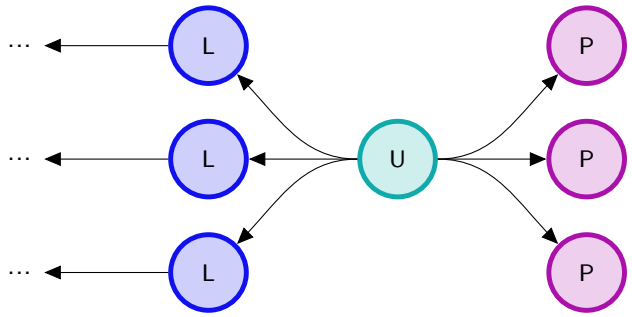


Figure 4.3: Example of one UserAccount contributing Locutions to multiple topics/threads, with multiple associated Personas

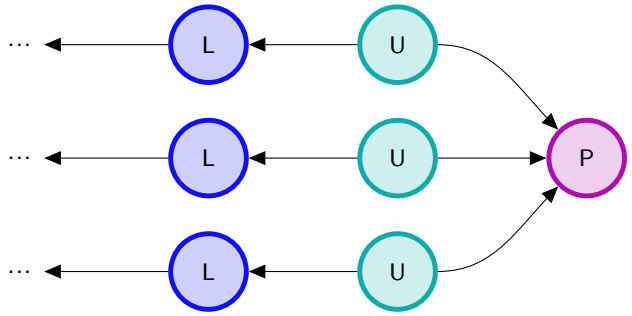


Figure 4.4: Example of modelling *"sock-puppeting"*: one Persona, linked to multiple UserAccounts

position, such that the position is rejected, but the negation of it is not inherently accepted) is carried to its conclusion, such that (certain) *ad hominem* attacks can undercut *all* of a participant's positions. However, without further additional changes, this method would require (for example) Personas to be attacked using standard Conflict nodes which, as a primarily logical construct, does not seem a suitable approach when modelling the notion of abuse. As a result, an additional pair of nodes was proposed, namely Personal-Conflict and Personal-Support, that represent rhetorical attacks and supports on a Persona. An example usage is shown in Figure 4.6. These nodes allow annotators to be specific, and differentiate between rhetorical and logical contributions when modelling.

Alongside this, while not the primary focus of this body of work, the topic of modelling social voting, reputation, and reaction systems has also been considered. These types of systems make up a key aspect of non-verbal argumentation on the social web, allowing users to show agreement or disagreement to a position, sometimes anonymously, without the need to articulate their own position, as well as potentially serving as a reputational currency. Users on Twitter, for example, are sometimes derided for being "ratio-ed" (receiving a high number of—usually hostile or argumentative—replies, compared to a low number of likes or retweets) for a poor argument.

Figure 4.7 shows one approach; namely, modelling each vote as a separate Locution, linking to an RA-node that logically supports (or a CA-node that logically attacks) the voted-on post.

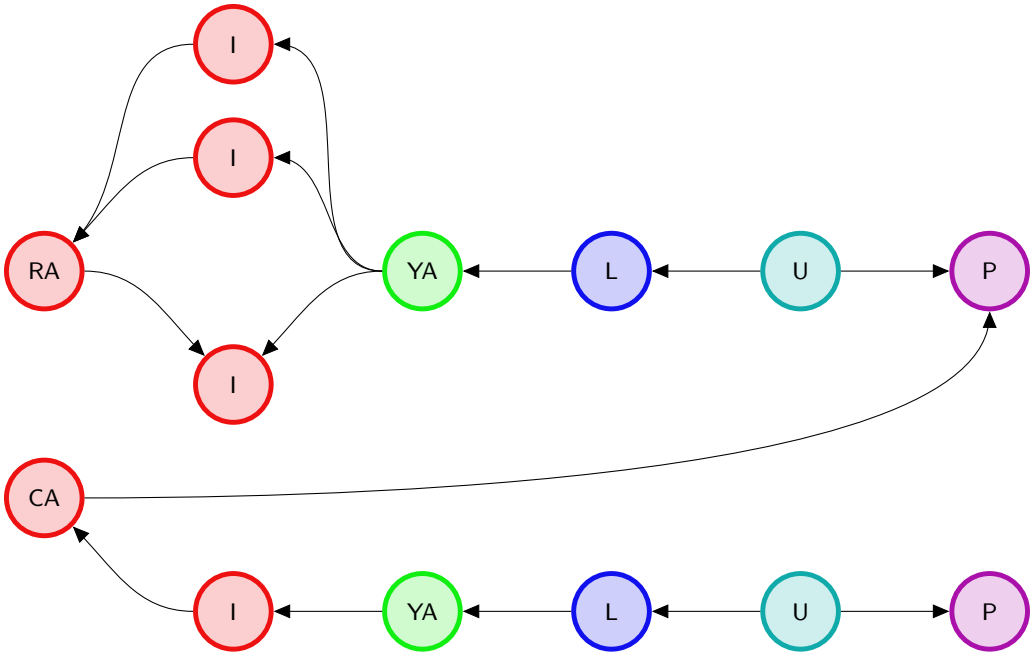


Figure 4.5: Proposal for representing abusive attacks, extending that shown in Figure 4.2 with the addition of Persona nodes

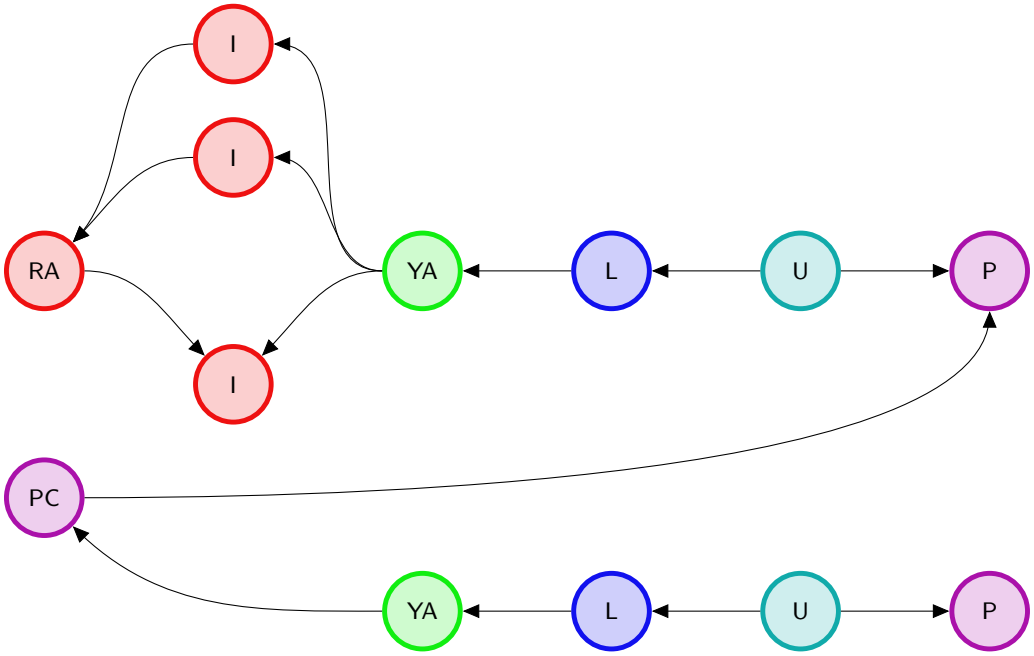


Figure 4.6: Proposal for representing abusive attacks, extending that shown in Figure 4.5 with the addition of Personal Conflict node

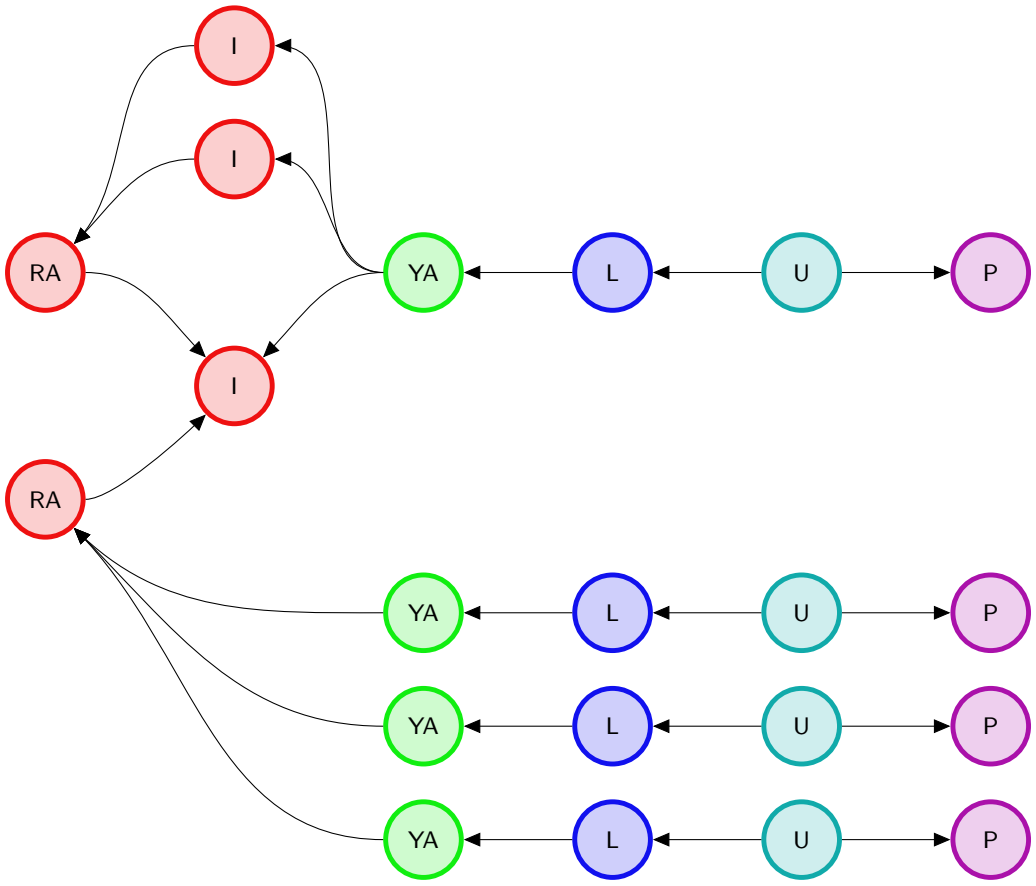


Figure 4.7: Modelling social reputation and reaction systems as logically supporting or attacking a particular idea

Alternatively, Figure 4.8 shows an approach in which votes are modelled as primarily rhetorical support (or attack) for the Persona *presenting* the idea, rather than the idea itself. The difficulty here is, of course, in that it is usually extremely difficult (when faced with a lack of contextual information) for the annotator to know the precise intent behind the reason a user either up- or downvotes a given Post. As a result, ASWO is capable of modelling both approaches, and allows annotators to choose the approach they feel is most appropriate; this should become easier in future, as some social platforms are beginning to increase the semantic information present in their voting and reaction systems. Facebook, for example, now allows users to react to a post with sad, shocked, or angry emoji, as well as the almost ubiquitous thumbs-up, to convey other emotions such as sympathy.

TA-nodes share a common theme with the *has_reply* relation of SIOC in that they are used to make links between locutions; however, in accordance with IAT, they can also be used to anchor certain inferences made by participants, most commonly interrogatives (for example asking for further information or evidence for claims that have been made), or those that make a counter-claim. These transitions are used when a Locution contributes to the argument without necessarily providing any direct information, but instead helps move the discussion to the "next stage", usually by asking questions or prompting further debate. Note that these transitions don't

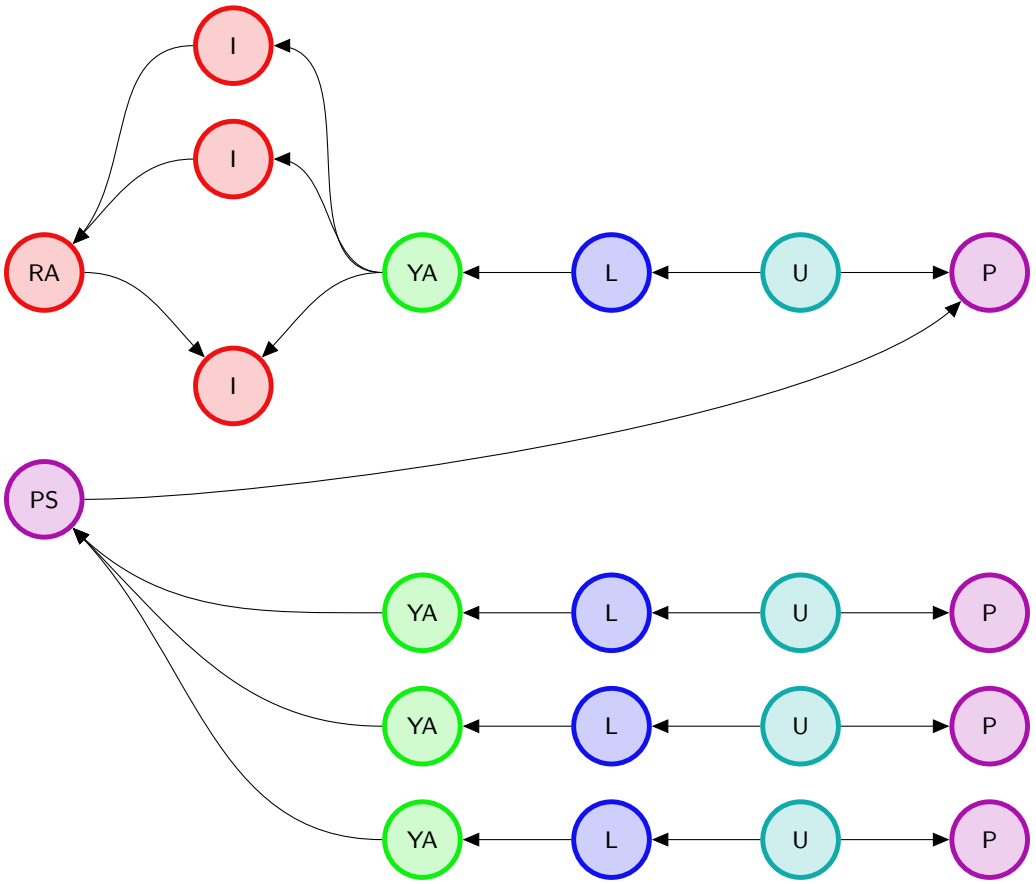


Figure 4.8: Modelling social reputation and reaction systems as rhetorically supporting or attacking the persona behind a particular idea

Table 4.1: Description of nodes added to the model, and their subsequent appearance in diagrams

Name	Description	Node
P-node	**Persona** nodes denote a person's social "character" that they assume during a discussion	P
PS-, PC-nodes	**Personal Support** and **Personal Conflict** nodes support/attack **Personas** in a non-logical way, implicitly undermining or upholding all their contributions	PS PC

necessarily move the discussion forwards, but can also be used to take the argument around in circles by asking questions in bad faith (or *sealioning*[1]).

These new nodes are summarised in Table 4.1, and the overall structure in Figure 4.9

[1] A term that originated with this Wondermark comic strip: http://wondermark.com/1k62/

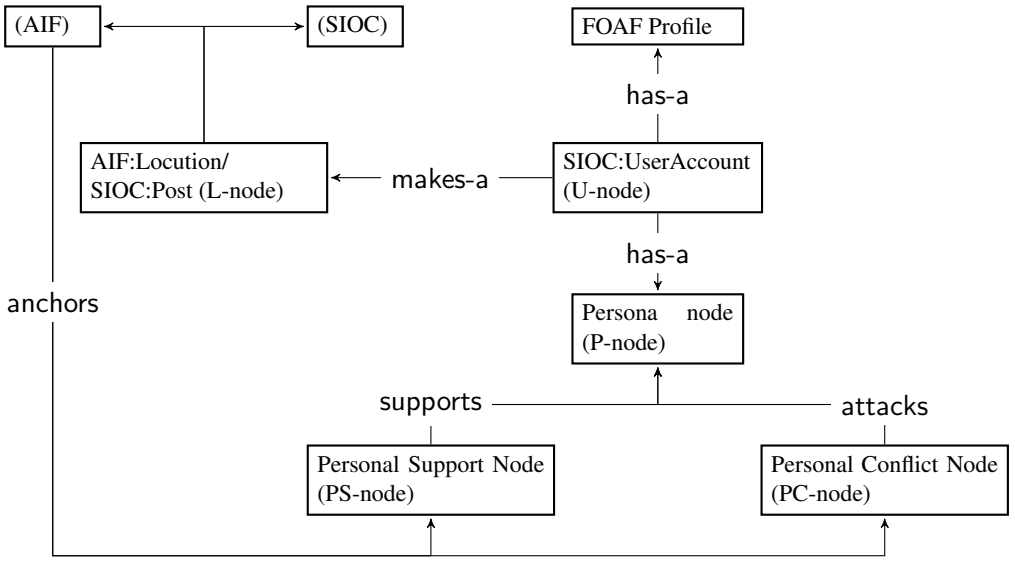


Figure 4.9: An overview of the ASWO structure

4.2 Experimental Application

To determine whether these additional nodes could be used to more thoroughly model social argumentation, ASWO and the augmentations made to the AIF and SIOC ontologies were trialled in an experiment to study the application of logical versus rhetorical techniques in eristic dialogue on the social web. As before, this investigation focused on three different areas of the social web, but used a larger sample size than previously: in total, 270 posts were collected and annotated. These were used to analyse the proportion of rhetorical contributions throughout the argument, analyse the relation between logical and rhetorical arguments used, and compare the features of the annotation structure with the content of each post.

4.2.1 Methodology

4.2.1.1 Data Collection

This investigation examines the same topic as as used in Chapter 3 (the U.S. government shutdown of 2013) but expands the amount of posts annotated to a total of 270 (90 from each social media platform).

During the course of this work, the Google YouTube API v2.0 was deprecated before the API v3.0 fully supported the retrieval of explicit replies to comments. Due to the importance of the ability to capture replies, the decision was made to use an alternative platform in this case study. To this end, YouTube was replaced with the social news and networking site Reddit. Reddit has a variety of topic-specific boards or "subreddits" that allow users to post to a collaborative pool of information; posts can then be up-voted or down-voted to show interest and/or accuracy.

Table 4.2: Metrics of discussions sampled from Twitter, Facebook and Reddit

Metric	Twitter	Facebook	Reddit	Total
Posts	90	90	90	270
Direct replies	77	0	67	144
Number of users	26	85	43	154
Average posts per user	3.5	1.1	2.1	1.8
Average words per post	15.83	41.36	42.34	33.18
Average characters per post	96.51	265.27	243.31	201.70
Time between first and last posts	0d 6h 53m 40s	3d 4h 51m 27s	3d 0h 50m 12s	n/a
Mean time between posts	04m 39s	51m 49s	49m 06s	35m 11s

Obama's official account on Reddit was inactive over the period of the shutdown; however, another user (unaffiliated in any official capacity with Obama) posted a link to Obama's official website (managed by Organizing for Action) to Reddit's politics subreddit[2] on 15th October 2013 (the same date as the official posts to Twitter and Facebook). The post reads *"Tea Party Republicans in the House of Representatives have already shut down the government because they couldn't derail Obamacare. Now they're threatening to cause an economic shutdown"*.

This thread was used alongside the previously acquired threads from Twitter and Facebook described in Section 3.1.2. Each UserAccount involved in the three threads was automatically designated a single Persona, as only one topic was monitored. This could be expanded if the same UserAccount took part in multiple threads on multiple topics, for example.

As with the preliminary work, forest fire sampling of the graphs was undertaken to provide a representative sample of the arguments that was feasible to annotate manually. For this investigation a larger sample size of ninety posts was used from within each discussion. Table 4.2 shows an overview of the sample structures and some key characteristics of each thread.

4.2.1.2 Annotation Method

With the changes to the ontologies in use (such as the decision regarding TA-nodes discussed in Section 4.1), and a larger amount of data needing to be annotated, the annotation method itself needed to be properly formalised to solidify reproducibility and minimise subjectiveness. Posts are annotated according to the scheme below.

Each post is considered to contain zero or more separate arguments. A YA-node is created for each argument made in a single post, and links the L-node to each I-node in the argument. Repeated information does not create a new I-node; instead the YA-node links to the I-node already present. All participants are assumed to have some implicit knowledge about the world in general and the topic at hand. This is to avoid the inclusions of trivial I-nodes that state information such as *"Barack Obama is president of the United States"*, or even *"Barack Obama is a person"*. Any information explicitly contained in a post that is deemed to be not in this set

[2] http://reddit.com/r/politics/1oij25

and relevant to the discussion at hand was included as an I-node. Information that meets one (or more) of the following criteria is not considered relevant:

- Off topic: posts that do not relate to the topic being discussed are not considered relevant. Example: *"Ataturk did revolution ! building moderate muslim network is oxymoron which has been destroy secular , democratic, rule of law in Turkey."*
- Conversational: similar to off-topic posts, those that are conversational in nature are not annotated as information-containing. Example: *"I thank you, have a good night!"*
- Meta-argumentation: while argumentation about how to argue "properly" is an interesting construct in itself, and an important aspect of rhetorical and eristic argumentation, it was considered out of scope for this particular study. Example: *"Down voting = disagree Upvoting = agree" "The rules say explicitly not to do that....."*

Support and attack between different I-nodes is denoted as described above: logical support through the use of RA-nodes, attack through the use of CA-nodes and preference with PA-nodes, while rhetorical support and attack utilises the new PS- and PC-nodes.

Some nodes in the resulting argument graph may not be complete, or lack edges, as a result of the nature of sampling used to reduce the social media thread. For example, it may be possible to detect that a user attacks another user's persona through implicit context of the locution (an abusive reply, for example), but not possible to determine exactly which user they are attacking.

4.2.2 Results and Analysis

4.2.2.1 Annotations

Table 4.3 shows an overview of the number of AIF and ASWO nodes added during the annotation process. Each platform has 90 L-nodes (one for each post), but the overall information content differs substantially, with Facebook and Reddit each having almost double the information content of Twitter. Twitter, however, had substantially more TA-nodes and PC-nodes than the other two platforms, suggesting both more back-and-forth debate and more aggressive, heated debate.

4.2.2.2 Argumentation Tactics Over Time

The way in which the argumentation structure changes and grows over time, in both a logical and rhetorical capacity, is presented in Figures 4.10-4.12, by graphing how the total number of logical support and attack nodes (i.e. RA- and CA-nodes) and rhetorical support and attack nodes (i.e. PS- and PC-nodes) changes with each post contributed to the overal argument thread.

Table 4.3: Summary of AIF and ASWO nodes found in annotated discussions collected from Twitter, Facebook and Reddit

Metric	Twitter	Facebook	Reddit	Total
L-nodes	90	90	90	270
U-nodes	26	85	43	154
TA-nodes	52	9	15	76
YA-nodes	58	74	70	202
I-nodes	56	98	86	240
RA-nodes	13	20	24	57
CA-nodes	18	1	34	53
PA-nodes	4	4	2	10
PS-nodes	2	2	3	7
PC-nodes	26	6	12	44
L- to I-node Ratio	45:28	45:49	45:43	9:8

As can be seen from the samples taken from Facebook (Figure 4.11) and Reddit (Figure 4.12), the use of rhetorical tactics in both of these case studies rises slowly compared to the use of logical tactics. However, within the sample taken from Twitter (Figure 4.10), rhetorical attacks rise more or less in parallel to the logical contributions.

In both samples from Twitter and Reddit, the distribution of logical supports and attacks also remain approximately equal. Due to the tendency of RA-nodes to be used for logical support within an argument, and the tendency of CA-nodes to be used between arguments, this highlights a greater engagement between participants within these debates than the Facebook sample, which has only one CA-node and comparatively much fewer instances of logical or rhetorical contribution overall. In all three examples however, rhetorical conflict far outweighs rhetorical support.

Overall, it appears that there is no sudden shift in tactics from arguing logically to adopting a rhetorical approach—rhetorical argument forms an underlying and consistent strategy throughout the thread of debate.

4.2.2.3 Argumentation Tactics per User

The proportion of logical versus rhetorical contributions made by each user was also examined, and is shown in Figures 4.13-4.15. These graphs show the contributions made by each user (ordered by total contributions overall). Looking at the samples taken from Twitter and Reddit (Figures 4.13 and 4.15 respectively), it appears that users made more individual contributions to the argumentation structure than those in the Facebook sample (Figure 4.14). This, along with the data in Table 4.2, also supports the suggestion that there is more engagement within these communities than those present in the Facebook sample.

All samples also display a tendency for rhetorical contributions to be distributed across the scale, with (weak) grouping towards either end. This implies that the users most likely to employ

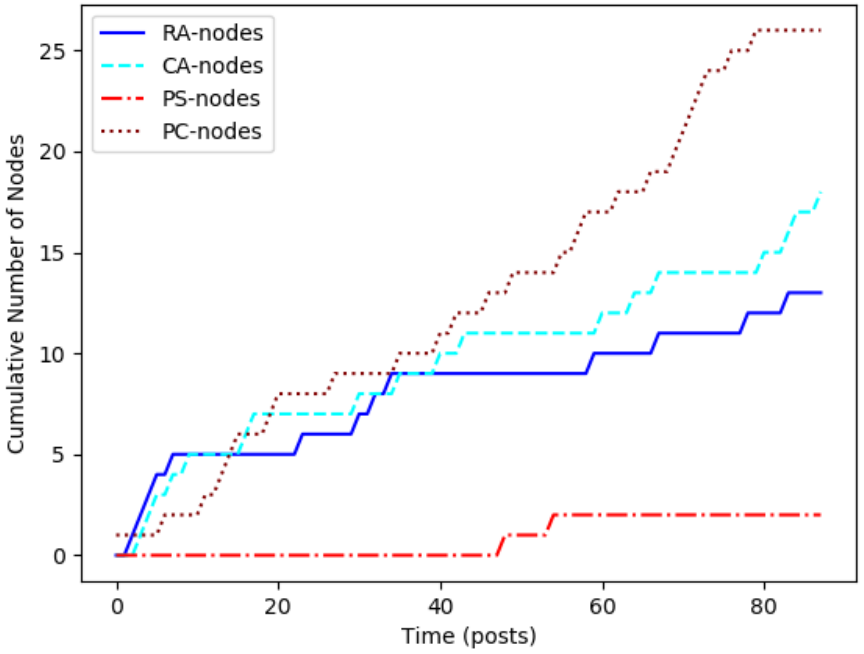


Figure 4.10: Cumulative use of logical and rhetoric tactics over time on Twitter

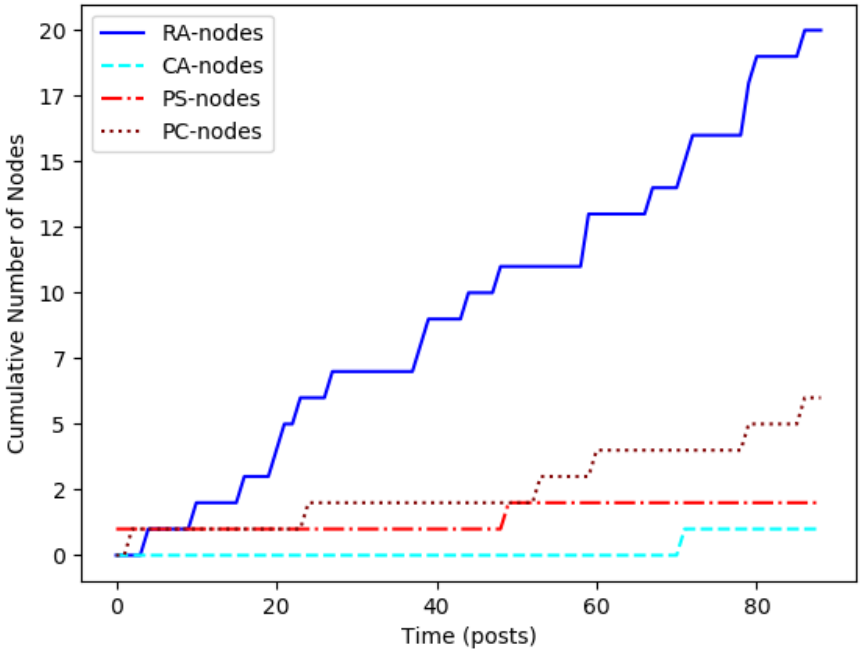


Figure 4.11: Cumulative use of logical and rhetoric tactics over time on Facebook

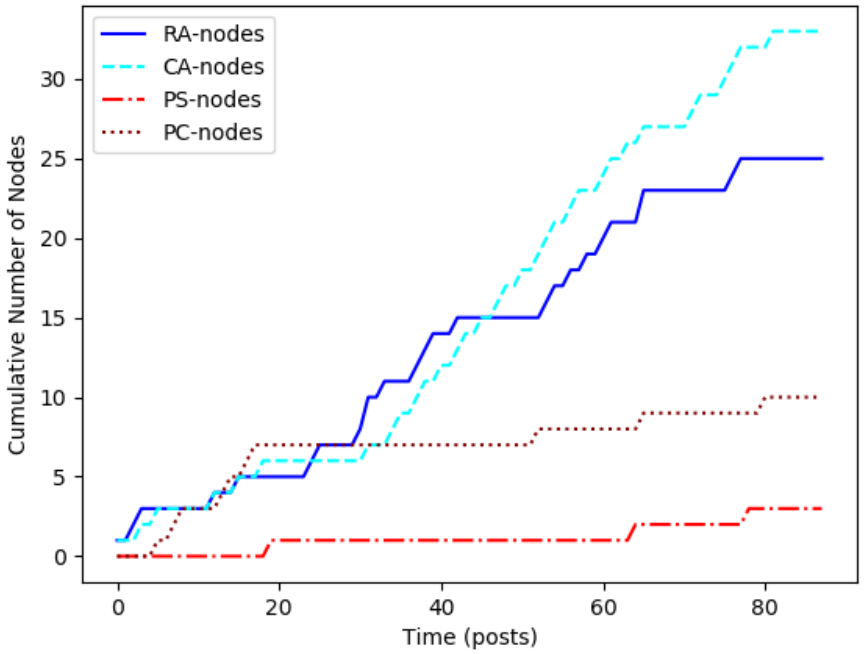


Figure 4.12: Cumulative use of logical and rhetoric tactics over time on Reddit

rhetorical techniques are those that contribute the most posts to the discussion overall, and those that make no logical contributions at all.

4.2.2.4 Correlation Between Argumentation Structure and Post Features

Pearson's *r* was used to test correlations between the structure of the annotated argument graph, including elements such as the number of logical or rhetorical supports or conflicts and replies to and from each post, and features of the post content and structure, such as post length, number of expletives, percentage of spelling errors and again, replies to and from the post. Replies in particular were viewed from both sides: that is, to analyse whether certain types of posts were more likely to be made in reply, or whether posts that were made in reply tended to contribute similar argumentation structures.

Due to the largely discrete (and often binary) nature of the features and values studied (the majority of posts, for example, are likely to contain either zero or one logical or rhetorical conflict nodes) the correlations are relatively weak, as can be seen when visualised (Figure 4.16). However, some notable correlations are presented in Table 4.4. These show potential early indicators of the structure and value of an argument. For example, as might be expected, longer posts are more likely to have greater contributions to the discussion. Posts that use a large number of expletives are likewise more likely to contain a rhetorical attack. When examining all three case studies together, posts made in reply correlated with posts that were replied to,

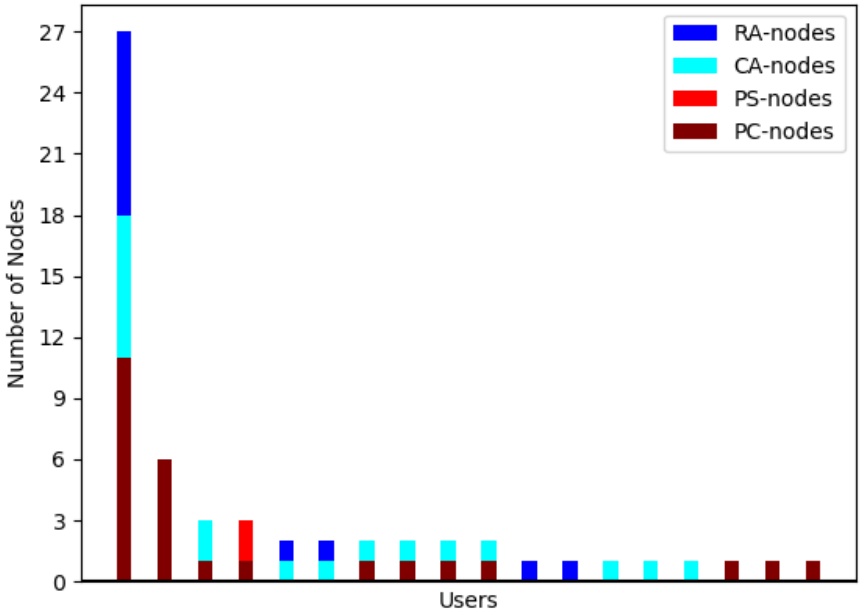


Figure 4.13: Logical and rhetorical contributions per sampled user on Twitter

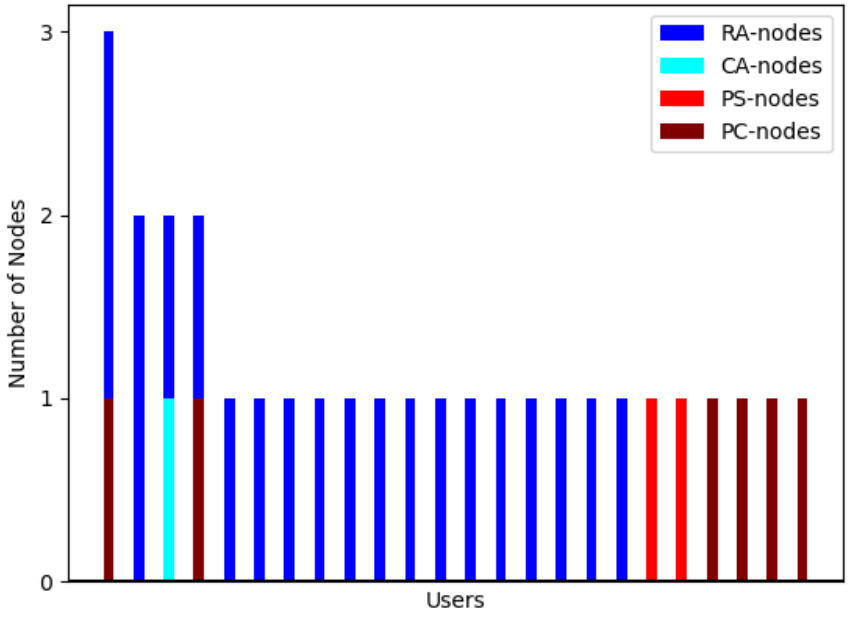


Figure 4.14: Logical and rhetorical contributions per sampled user on Facebook

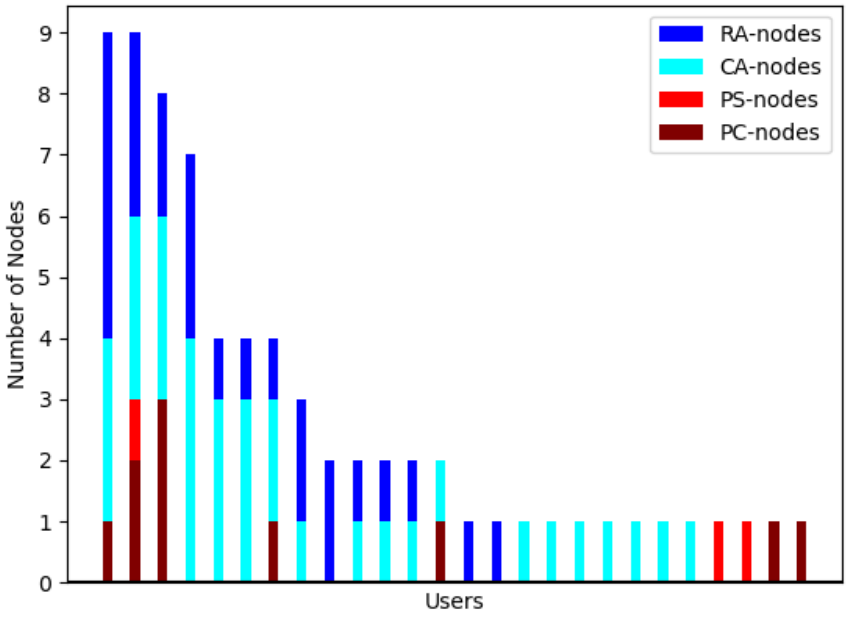


Figure 4.15: Logical and rhetorical contributions per sampled user on Reddit

implying that when one or more users engage in a discussion, they are more likely to be engaged with in return.

4.3 Further Proposals

Following on from the investigations in Section 4.2, further additions to the ASWO were made to capture greater depth of relationships between socio-rhetorical arguments.

Two new nodes, Faction and Audience nodes, were introduced to represent abstract groupings of Personas; a Faction is any grouping of Personas and can potentially include those outside the Thread. For example, a Faction could denote a political or religious group, members of a social media community, or simple those participants for or against a particular position in the debate. Whereas, the Audience node represents (again, abstractly) all Personas currently participating in, or observing, the discussion.

Another introduction was the Implication node, which allow analysts to represent a participant implying a particular relationship between two (or more) nodes, such as Personas or the aforementioned group nodes. These can be combined with the Personal Support/Conflict nodes to indicate whether the implication is positive or negative. For example, a participant could imply that their opponent is a part of an undesirable group to undermine their authority.

Table 4.4: Notable correlations between structural argumentation annotations and post features

Case Study	**Argumentation Structure**	**Post Feature**	**Pearson's r**	**(p-value)**
Twitter	Personal attacks	Number of replies to this post	0.325	0.002
Twitter	Personal attacks	Percentage of spelling errors	0.301	0.004
Twitter	Personal attacks	Expletives	0.462	< 0.001
Facebook	Original premises	Reputation ("Likes")	0.332	0.001
Facebook	Original conclusions	Reputation ("Likes")	0.329	0.002
Facebook	Logical inferences	Reputation ("Likes")	0.343	0.001
Facebook	Logical conflicts	Emoticons	0.500	< 0.001
Facebook	Logical conflicts	Expletives	0.397	< 0.001
Reddit	Original premises	Post length	0.335	0.001
Reddit	Original conclusions	Post length	0.333	0.001
Reddit	Logical inferences	Post length	0.476	< 0.001
Reddit	Logical conflicts	Number of posts replied to	0.435	< 0.001
Overall	Number of replies to this post	Number of posts replied to	0.417	< 0.001

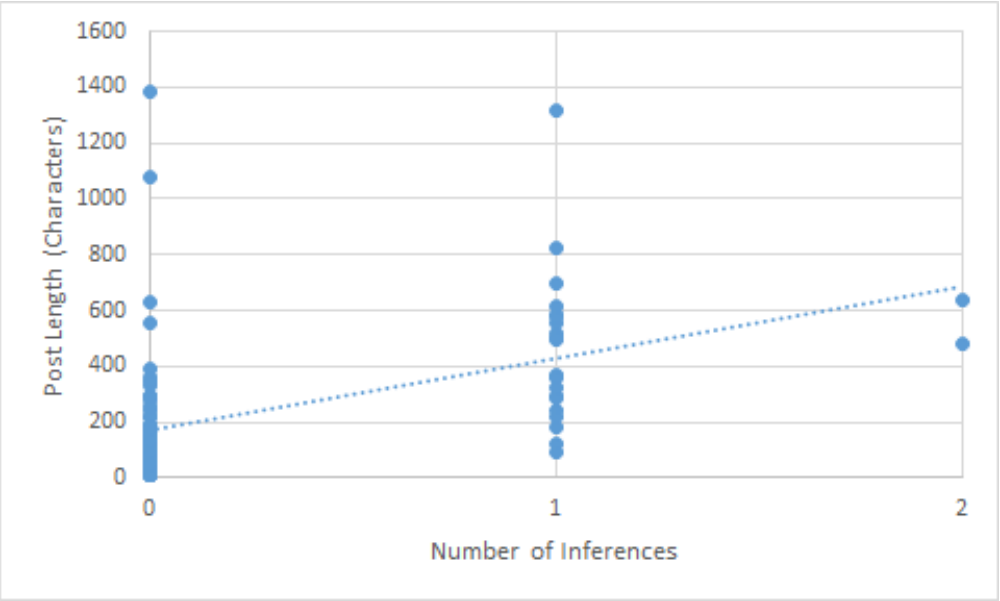


Figure 4.16: Post length correlated against number of logical inferences, on Reddit. Pearson's $r = 0.476$ ($p < 0.001$)

Table 4.5: Description of further nodes added to the model, and their subsequent appearance in diagrams

Name	Description	Node
F- and A-nodes	**Faction** and **Audience** nodes represent groups of personas	F A
Im-node	**Implication** nodes indicate a relationship that the participants can not be sure exists	Im

These additions aim to provide atomic "building-blocks" that can be reused to model a wide range of social, rhetorical and "extra-logical" aspects of argumentation. Here, these rhetorical nodes are used to show how some examples of common logical fallacies, with particular focus on the rhetorical force behind them, can be modelled under ASWO.

4.3.1 *Ad hominem*

Ad hominem ("to the man") arguments attack a person's character, without attacking their argument. However, they can be a viable tactic in rhetorical debate and can introduce both new I-, CA- and PC-nodes to the structure when modelled. Figure 4.17 shows how a fictitious example of a reasonable *ad hominem* argument (Walton, 1987) *"You don't have any qualifications in that area, don't make such broad statements"* can be modelled. Figure 4.18 shows a model of a more aggressive variant of the tactic that disparages someone's argument and them as a person: *"They're an idiot, don't listen to them."* Figure 4.19 shows an example model of an abusive argument that contains no information, instead attacking the person directly and trying to shut

them out of the debate: *"**** off and die!"* These examples in particular show that a fallacy can take multiple forms and have multiple logical and/or rhetorical contributions to the overall discussion.

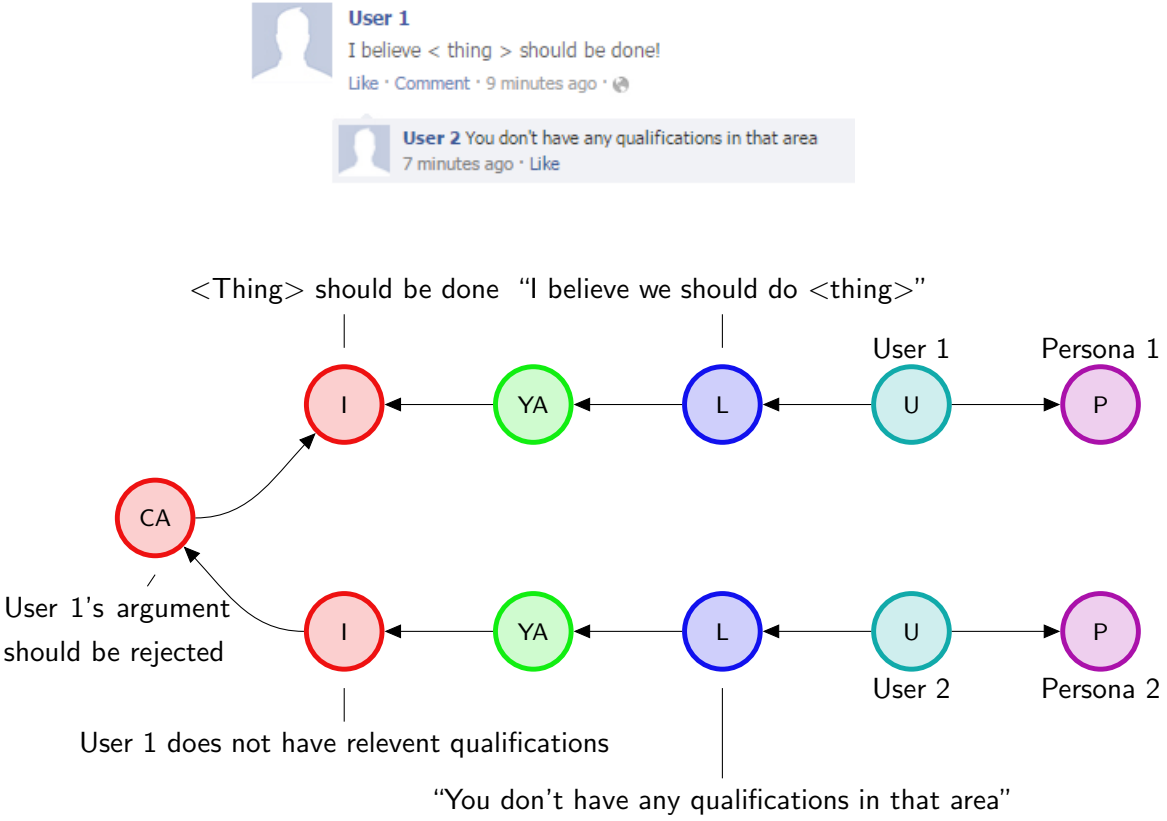


Figure 4.17: Example of a reasonable *ad hominem* attack

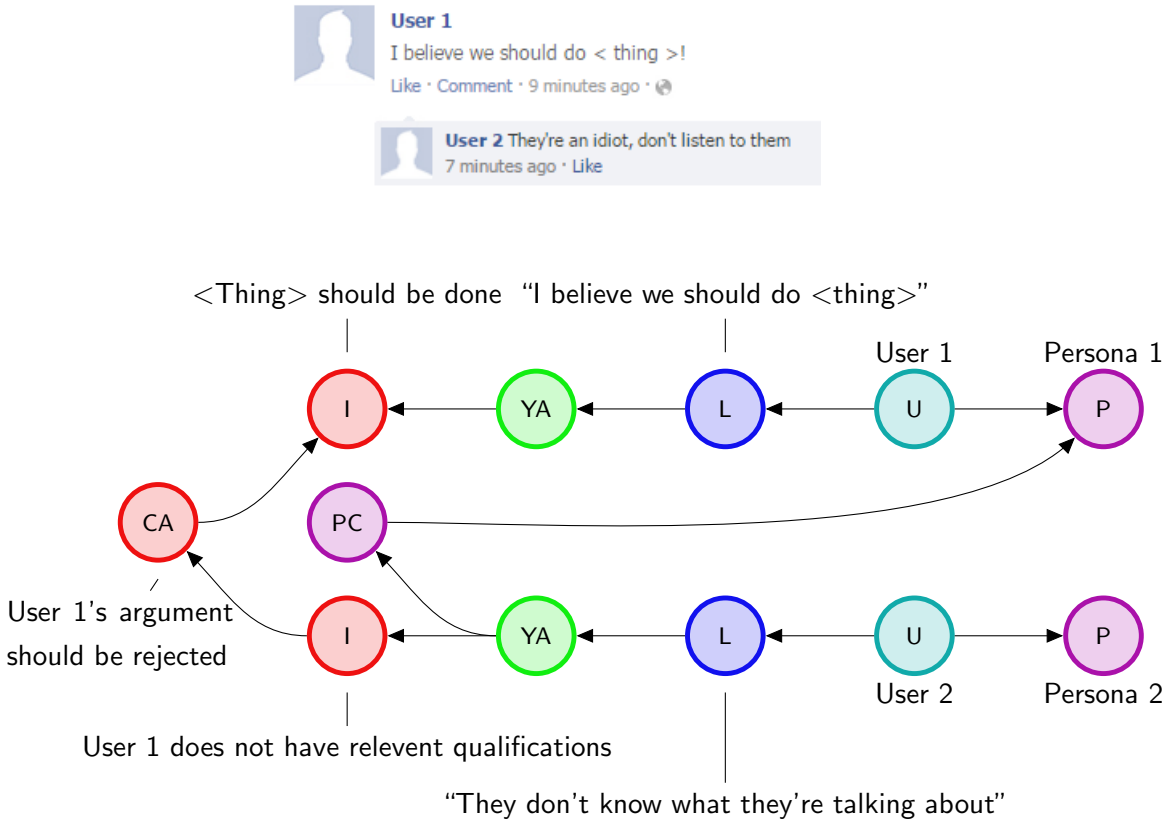


Figure 4.18: Example of an *ad hominem* attacking both persona and argument

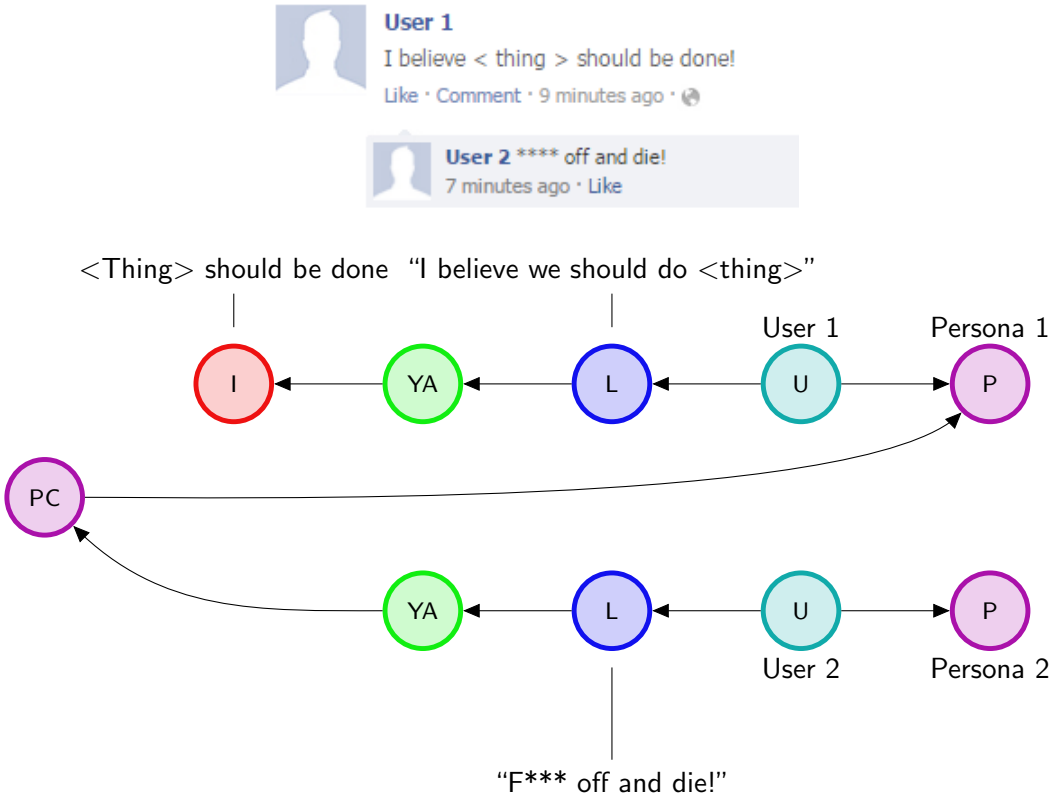


Figure 4.19: Example of an abusive *ad hominem* attack

4.3.2 Appeal to Consensus

The appeal to consensus is the fallacy that because a claim is popular or widely-held, it is true. An example of this can be shown in Figure 4.20.

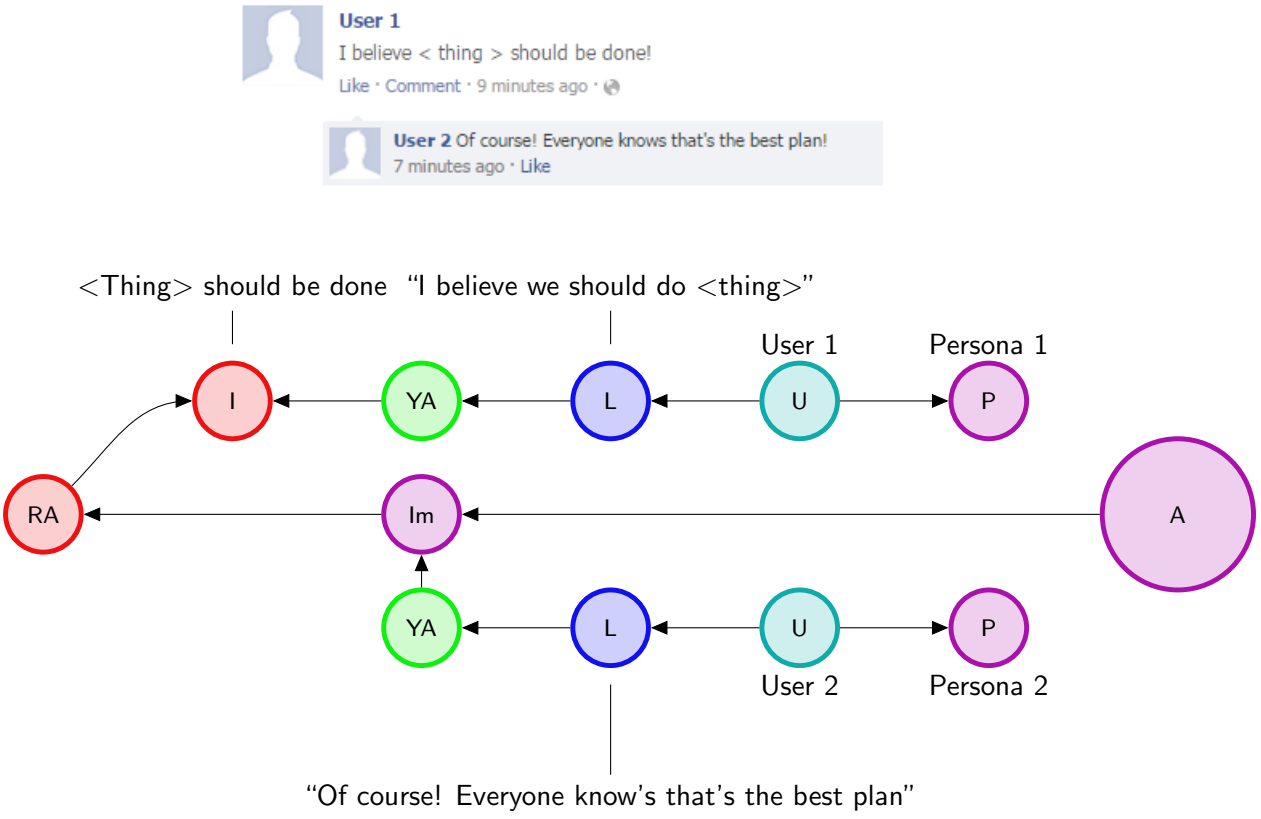


Figure 4.20: Example of an appeal to consensus

4.3.3 Association Fallacy

The association fallacy is the notion that because a person is associated with, or shares the views of, an undesirable group, their claims are wrong. An example of this can be shown in Figure 4.21.

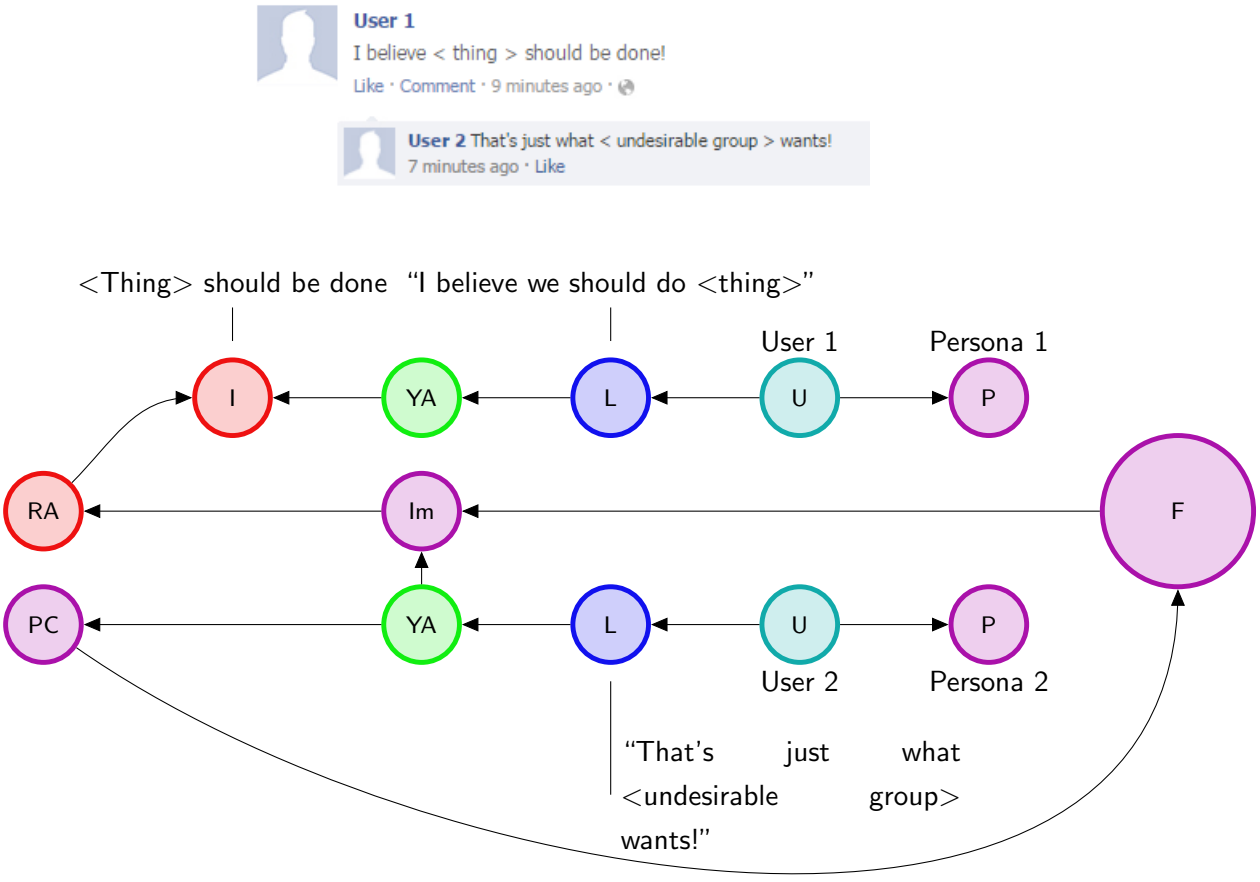


Figure 4.21: Example of the association fallacy

4.3.4 Appeal to Humour

An appeal to humour is a technique by which a participant in the debate attempts to ingratiate themselves with their audience by making a joke about the situation as shown in Figure 4.22. This can be coupled with an *ad hominem* attack, when the joke is made at someone else's expense, as shown in Figure 4.23.

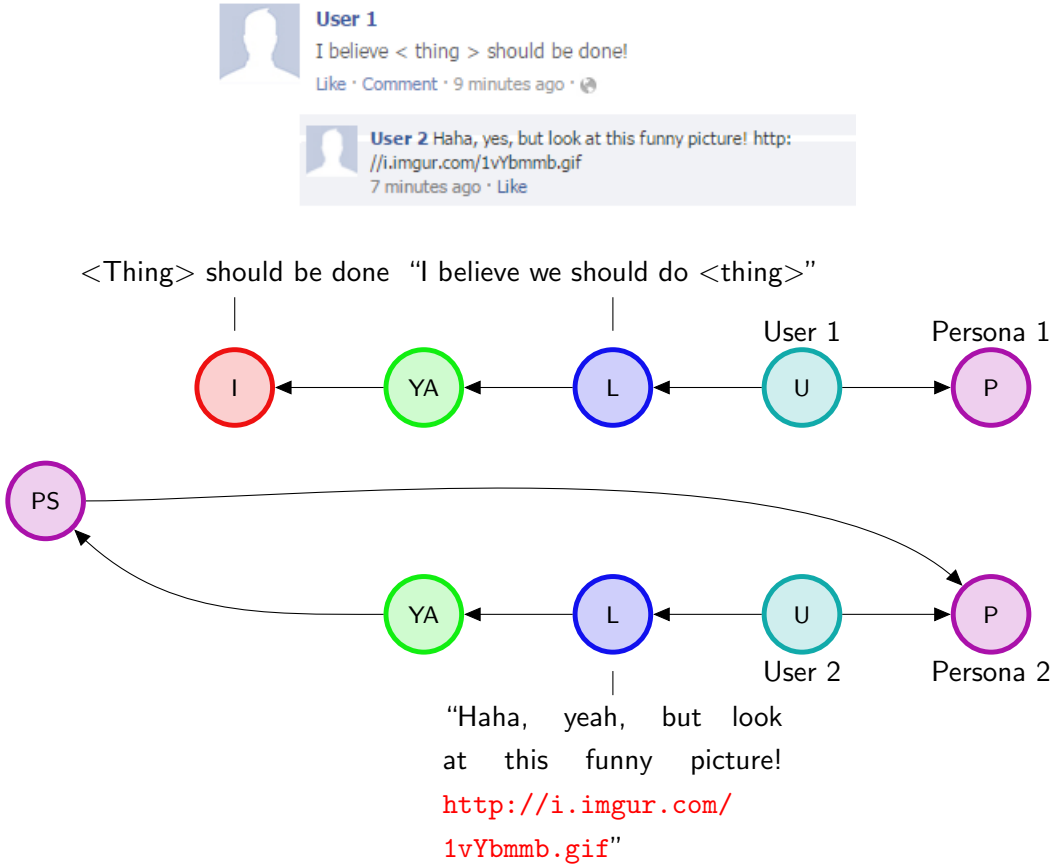


Figure 4.22: Example of an appeal to humour

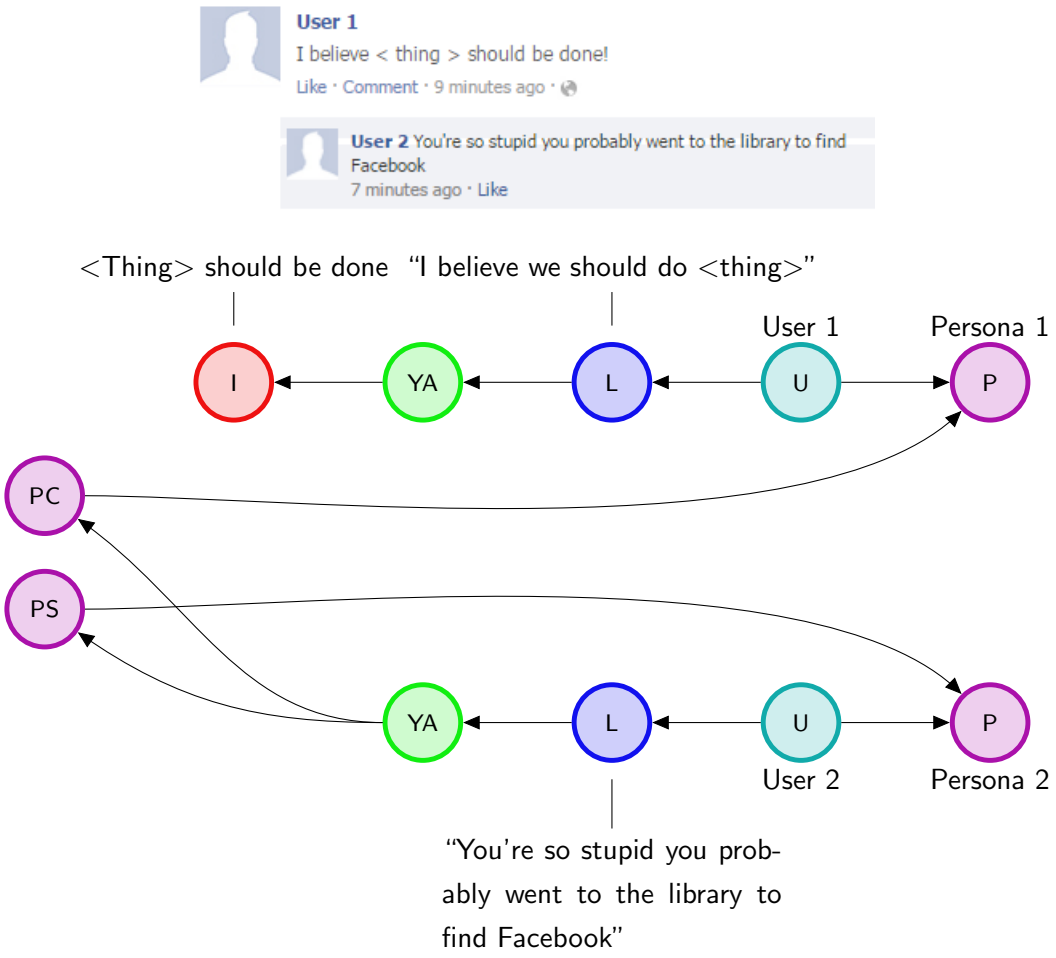


Figure 4.23: Example of an *ad hominem* appeal to humour

4.4 Expert Review

Six experts, from the fields of argumentation systems, web science, philosophy, and linked data, were chosen to review these proposed additions to the model. Ethical approval for this study was granted under code 17924; the application is detailed in Appendix A.1.

Experts A and B have a background in argumentations systems and modelling argumentation, and are familiar with the AIF. Expert A is a computer science lecturer whose research is concerned with argumentation-based models of communication and formal reasoning, with interests in AI and behaviour change. Expert B is a post-doctoral researcher with degrees in library and information science, mathematics, and liberal arts whose thesis focused on the problem of analysing, integrating, and reconciling information in online discussions.

Expert C is a web-science graduate student, researching the relation between social structures in virtual worlds and the real world, with a focus on practices of gender and power.

1. **User 1:** *The tech industry is often biased against women*
   User 2: *@User1 You would say that, you're a woman*
   User 3: *@User1 **** off and die you ****ing nazi before I come and **** you up*

2. **User 1:** *Guns killed 33,000 people last year, they need to be banned*
   User 2: *@User1 And a lot of those were minors*
   User 3: *@User2 According to who?*

3. **User 1:** *What does Barack Obama call illegal aliens? Undocumented democrats!*
   User 2: *@User1 You're so stupid you probably went to the library to find Facebook*

Figure 4.24: The three argumentation samples the experts were asked to model

Expert D is a philosophy graduate student, specialising in ethics, moral obligations and with a background in argumentation and formal logic.

Experts E and F are specialists in the area of open and linked data working in web and data innovation and development. Expert E is an institutional open data specialist and Expert F is a senior technical specialist.

4.4.1 Methodology

Each expert was provided with a document describing the background of this area and an overview of the existing models (reproduced in Appendix B). They were then asked to model three argumentation samples shown in Figure 4.24, illustrating a variety of different rhetorical structures, by speaking aloud and/or sketching with pen and paper. They were then shown the additions to the model, and asked to model the three argumentation samples again. They were then asked a series of semi-structured question aimed to evaluate their thoughts on how best (and whether) to model social (and anti-social) argumentation, the completeness of the ontology, the clarity of the ontology and the consistency of the ontology.

4.4.2 Results and Analysis

This section gives an overview of the key points discussed by the experts along the themes of modelling social argumentation, completeness, clarity and consistency (and relevant sub-themes).

4.4.2.1 Social Argumentation

Value

Each of the experts agreed that there was value in modelling social argumentation, Expert F going so far as to say they believed there was no argument that didn't have social components.

> (A) *"...if we're going to have a realistic model of how people argue, we've got to look at how people really argue rather than how our "ideal reasoner" would argue"*

Expert D discussed how understanding the nuances of how people argue socially could lead to ways of helping or encouraging them to argue in a more cooperative or polite manner.

> (D) *"I think modelling social argumentation is very important...I want to say it's useful in trying to help people argue "better"."*

Challenges

The challenges of modelling social argumentation the experts foresaw were mostly a question of scale. The sheer volume of data in a social media discussion can be overwhelming, particularly when considering the speed with which in can grow.

> (C) *"Even in quite a simple back-and-forth argument, there's quite a lot going on...scale is a challenge "*

In addition to this, experts also noted the variety of information, which is often contextual, such as references to current events, or cultural "in-jokes".

> (B) *"...enthymemes, humour, there's lots of missing information, there's lots of playing to particular audiences...there are lots of things that are current events or would only make sense to a particular group"*

Abuse and Threats

Experts A and D explained that, although abusive argumentation was a potentially valuable area to explore, they would not consider it as valid when modelling an argumentation structure (as they focused broadly on dialectic arguments and that was the current standard for their domain).

> (A) *"I, personally, tend to ignore all of those because I'm...focusing on the informal proof structures"*

Expert B explained that it depended very much on the purpose of the model — in some cases it may be important to model threatening and abusive attacks specifically so they can be excluded when presenting the model to users.

> (B) *"...it's hard to exclude them...if you think about what you're going to do with the model...do you want to retrieve threatening and abusive comments? Well you might want to exclude them from being retrieved, which also makes it relevant to model that"*

Expert E also noted that excluding this type of argument can lead to confusion if a particular abusive comment changes the course of the argument, or causes the quality of the rest of the discussion to degenerate.

4.4.2.2 Completeness

Implicit and Explicit Premises

Expert B noted that, as many annotations have the potential to be subjective, it would be possible to extend this to include further subjective annotations such as an analyst's confidence in a particular reading of an inference. Expert C had similar views and discussed including mappings of a participant's agreement or disagreement with key positions in the dialogue as well.

> (B) *"I think when people model arguments it's pretty common to infer the reading, and what's interesting is that there can be multiple readings. So it wouldn't be wrong to...put in some interpretation, as long as it's clear it's an interpretation and there can be others."*

Social Metadata

Experts A and B both made explicit mention of the ability to mark certain posts as being in direct response to other participants in the discussion as a useful addition to argumentation frameworks.

Expert F discussed the potential for an "activity" score for each locution, derived from the social meta-data of each post (e.g. number of replies, number of up- or down-votes or number of retweets); this metric could be derived on a per-purpose basis to allow analysts to correctly categorise different platforms for their own needs, and to highlight key areas of the discussion that had solicited or stimulated large amounts of discussion.

> (F) *"One other thing... is other people's opinions of statements. A lot of systems have thumbs up and thumbs down...what you need is, I think, an audience response"*

All experts agreed that to adequately model social argument that it was necessary to include further context about the participants, such as demographic information where available, such as by linking the SIOC UserAccount to a FOAF Agent, or additional information about key events related to the discussion to maintain relevance of the model for future analysis, and to limit the number of assumptions needed to be made by analysts.

4.4.2.3 Clarity

Ambiguity

Expert's B and D were concerned that, when faced as an analyst with a statement that appeared ambiguous (for example, a statement of support that could be interpreted as genuine or sarcastic) they may struggle to accurately and objectively model it. Expert B suggested it could be valuable to associate each annotation with a confidence level (determined by the annotator), whereas Expert D proposed a means of allowing analysts to mark such relations as existing without committing to associating them with either a support or an attack.

Generalisation

Expert F proposed a similar solution, by means of generalising the model to include super-classes of Support and Conflict. "Personal" conflict, for example, is perhaps too specific a name for all non-logical conflicts: there are rhetorical attacks that can target institutions or accounts run by software, but also, importantly, positions and information. These Support and Conflict super-classes would encompass Logical Support/Conflict and Rhetorical Support/Conflict and could then be further sub-classed to provide more specific instances of each, where apparent, allowing analysts to defer when unsure.

> (F) *"If anything I think maybe your default conflict is a superclass - everything is a conflict, and one of the subclasses is a...rational argument. But then you've also got personal attack, ad hominem...these are all alternatives to rational argument, but at the default it might be worth allowing modelling of a conflict. Not a conflict as it is in the original model, but as a superclass of interaction."*

4.4.2.4 Consistency

Internal Consistency

In terms of inter-rater reliability — whether two analysts attempting to model the same argument would reach the same result — the experts were much more divided. While they agreed that the objective parts of the model (i.e. the locutions, user account and, in most circumstances, the persona) could be modelled identically (and in most cases, automated), Experts C and B felt that both analysts would reach the same conclusion overall with minor deviations, whereas Experts A , D and E disagreed, stating there was too much subjective information to model identically. Expert A felt that the analyst would naturally perceive the argument through their own lens of cultural and social context and Expert D noted the different levels of detail an analyst may choose to use, whether focusing only on premises that have been explicitly stated, or including additional implicit information.

> (A) *"...rather than having the minimal number of nodes and encouraging people to just misuse them, I would rather say 'Here's a definite type of argumentation we want to capture and share..."'*
>
> (E) *"Whenever you try to model anything in a formalised system...if you give two people the same thing...unless it's something really simple, they will always find two different ways of modelling it"*

How important this is was also a matter of some debate: Experts B and C felt that it was likely there would (and should) be one "correct" representation of an argument. Experts D and F agreed to an extent, citing their proposals for handling ambiguous content being able to aid annotators in this regard, so that if the model could not be complete, it could be consistent. Expert A felt that ideally analysts should reach the same conclusion but in practice, the subjective nature of the task might make this impossible. Expert E felt the consistency of annotators would, in practice, be less important and would be a factor of the intended purpose of the model.

External Consistency

The majority of experts felt that the additions to the ASWO were consistent with the nodes used in the AIF. However, Experts C and F disagreed, suggesting that the ASWO was in fact *intentionally* inconsistent with the AIF, because they were developed for different purposes.

> (C) *"Consistent with [the AIF], maybe not, but building on? Definitely"*

4.5 Final Proposals

Based on the expert review, the following changes were made: Personal Support and Personal Conflict have been explicitly renamed to Rhetorical Support and Rhetorical Conflict, to clarify instances where it is not a *person* that is directly being targeted in this manner, but some other entity (encompassing situations where, for example, an institution or organisation can be thought of as having a Persona for the purpose of the debate). Also for the sake of clarity (and to avoid negative connotations), Faction was renamed to Group. To allow for greater flexibility of annotation, Implications were split into two types: Implied Relationship, to account for occurrences when participants specifically imply that a Persona belongs to a certain group (for good or ill); and Implied Belief, for when they imply that a Persona, Group, or the Audience itself, holds a specific belief. The full set of ASWO nodes is summarised in Table 4.6, and the structure in Figure 4.25.

4.6 Summary

In this chapter, a number of proposals aimed at further augmenting the AIF and SIOC to accurately model eristic social argumentation are described, termed the Argumentation on the Social

Table 4.6: Description of the final set of nodes used in ASWO

Name	Description	Node
P-node	**Persona** nodes denote a person's social "character" that they assume during a discussion	P
G- and A-nodes	**Group** and **Audience** nodes represent groups of personas	G A
RS-, RC-nodes	**Rhetorical Support** and **Rhetorical Conflict** nodes support/attack **Personas**, **Groups**, or the **Audience** itself, in a non-logical way, implicitly undermining or upholding all their contributions	RS RC
ImR- and ImB-nodes	**Relationship Implication** nodes indicate a relationship that the participants can not be sure exists; similarly, **Belief Implication** nodes indicate a belief	ImR ImB

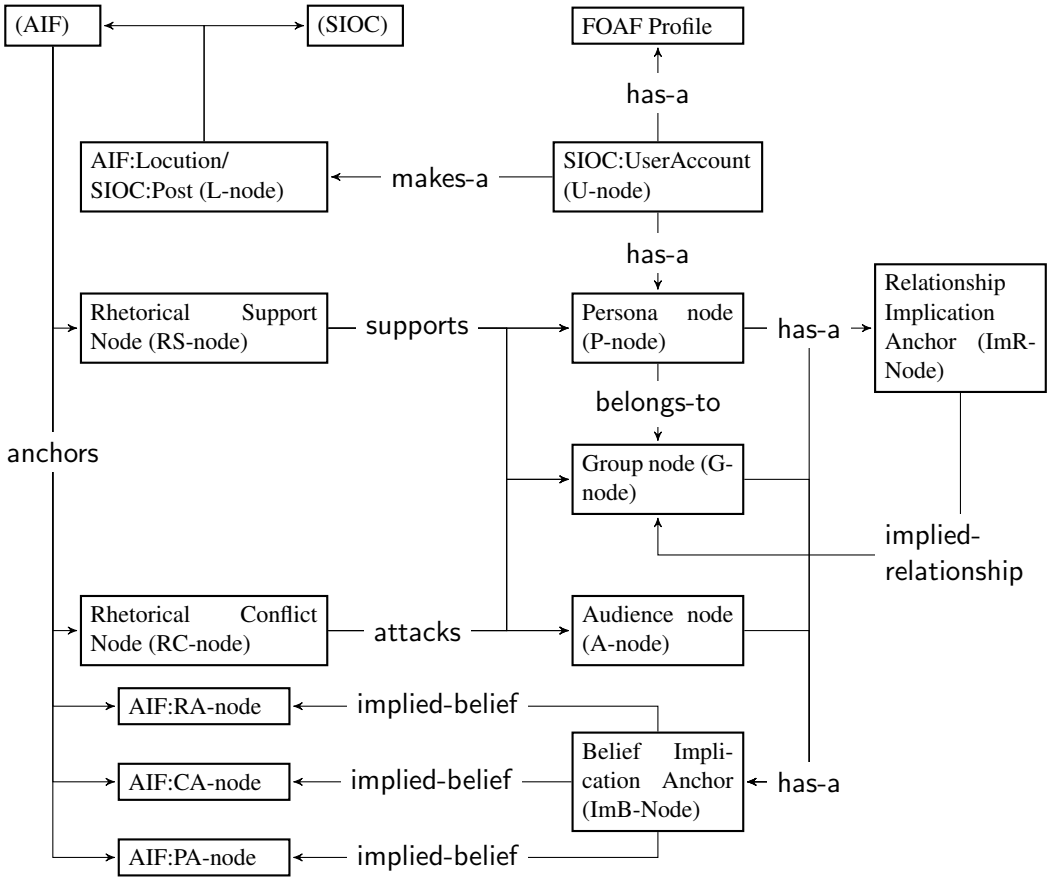


Figure 4.25: An overview of the final ASWO structure

Web Ontology, to provide an answer to research question $R3$ (*How should rhetorical techniques be included in a model of eristic argumentation on the social web?*). These proposals are trailed in an investigation of argument on a real-world example of social web discussions, and further proposals introduced as a result. These further extensions to the ASWO were evaluated through an expert review, which highlighted the key strengths of this model (as well as further highlighting the importance of being able to model eristic argumentation ($R1$)), such as the ability to model directed replies, the ability to model the audience of the discussion, and the ability to model instances of irrational and eristic argument that were previously difficult or impossible to achieve with the AIF alone ($R2$). These results were presented at ACM Hypertext (Blount et al., 2015a) and the International Conference on Computational Models of Argument (Blount et al., 2016).

Because social argumentation can rely heavily on nuanced contextual information (such as the ability to recognise humour, sarcasm or references to current events) it is likely impossible to model it in such a way that it could be automatically reasoned over. However, because the ASWO provides additional information about rhetorical tactics in use, human analysts can explore the resulting structure in greater detail and context. This can also potentially be used to highlight areas of particular interest, or assist in community decision-making environments.

The ASWO is next used to assist in constructing a much larger dataset of social argument, to re-examine the techniques and tactics used on three different social media platforms under this new context, for later use in an experiment to determine how logic and rhetoric affects an audiences perception of argumentation.

Chapter 5

Case Studies of Argument

To show that the Argumentation on the Social Web ontology is sufficient for modelling eristic argumentation such as that present on the social web (*R*3), a larger corpus of social media posts was annotated, and then examined in depth. Social media discussions were gathered from three different social platforms; these discussions were then annotated using a categorisation system based on the Argumentation on the Social Web Ontology proposed in Chapter 4. The distribution of social features (such as replies etc.) and annotations was examined and discussed, and a closer analysis (including more in-depth modelling) of three randomly selected threads was conducted to discuss some of the features that can be observed on different social platforms in different contexts.

5.1 Methodology

5.1.1 Data Sample

Social media posts for this experiment were sourced from Facebook, Twitter and Reddit. As with Chapters 3 and 4, to generate arguments and discussions of suitable quality, topics observed need to be both popular and polarising. News distribution networks were chosen as a data source of suitable topics of this nature. From Facebook and Twitter, posts were collected by gathering the 100 most recent posts from five major news distribution networks (*BBC News*, *Sky News*, *CNN News*, *The Guardian*, and *The Daily Mail*) and all replies associated with them. These particular news networks were chosen to be representative of popular English-language news sources, from both sides of the political spectrum. Because of the "subreddit" structure of Reddit (there being many "sub" sites focusing on individual topics), and because these institutions do not maintain official accounts or an otherwise high-profile, active presence on Reddit, this portion of the data sample was sourced instead by collecting the top 500 "hot" posts from the World

Table 5.1: Distribution of total number of comments in each thread

Platform	**Min.**	**Lower Quartile**	**Median**	**Upper Quartile**	**Max.**	**Mean**	σ
Facebook	3.00	43.00	96.00	189.00	2478.00	179.976	274.449
Twitter	3.00	5.00	9.00	15.00	193.00	11.717	12.693
Reddit	3.00	4.00	10.00	70.00	4938.00	157.081	522.048

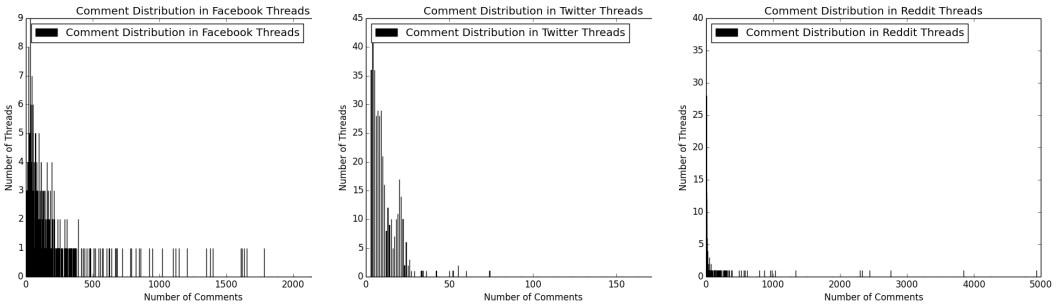


Figure 5.1: Number of comments per thread on Facebook, Twitter, and Reddit

News subreddit[1]). "Hot" posts, the default view of subreddits, are determined by those with the most upvotes, and are weighted towards more recently active posts (Van Mieghem, 2011).

These threads were then pruned to ensure they contained some form of discourse; that is to say they contained at least two replies, made by at least two different users. Applying this filter resulted in 500 threads from Facebook, 403 from Twitter and 247 from Reddit. To remove outliers from this sample (and ensure that the gathered threads were more representative of each platform) these threads were then examined based on a number of different factors to determine the shape of a "typical" thread for each platform; any thread not falling within the interquartile range of each of these factors was removed from the sample. These factors were the *total comments*, *comment length*, *comments per user* and *number of (internal) replies*. Tables 5.1, 5.2, 5.3, and 5.4 show a further breakdown of each of the measures used to determine a typical thread for this purpose, and Figures 5.1, 5.2, 5.3, and 5.4 show a comparison of the respective distributions. Larger versions of these graphs are also shown in Appendix C.

Table 5.1 (and Figure 5.1) shows the distribution of the total number of comments in each thread. Naturally, the minimum number of comments for each thread is three; the original post, and two additional, the qualify it as discursive. However, the maximum number of comments was substantially larger, and varied hugely between platforms. Twitter had the smallest range, with 193, and a correspondingly tighter spread. Both Facebook and Reddit had a much greater maximum, with approximately 2,500 and 4,000 respectively. However, in both cases, these were decidedly outliers.

Table 5.2 (and Figure 5.2) shows the distribution of the average comment length in each thread. In a similar pattern to the total number of comments, the length of comments on Facebook and Reddit also displays a distribution of a curve centred around the lower end of the range with

[1] https://www.reddit.com/r/worldnews/

Table 5.2: Distribution of comment length (in characters) in each thread

Platform	**Min.**	**Lower Quartile**	**Median**	**Upper Quartile**	**Max.**	**Mean**	σ
Facebook	24.00	68.00	103.50	161.25	715.00	127.936	90.983
Twitter	38.00	69.00	83.00	96.00	136.00	82.553	18.835
Reddit	27.00	97.00	173.00	260.50	697.00	200.870	132.320

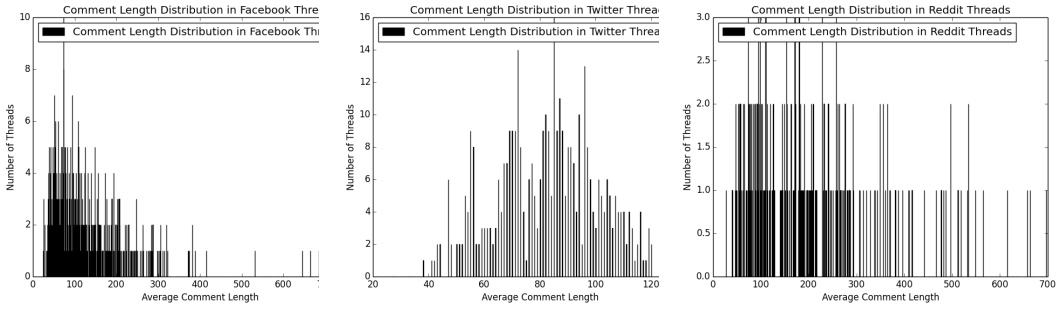


Figure 5.2: Comment length distribution on Facebook, Twitter, and Reddit

Table 5.3: Distribution of the number of comments per user in each thread

Platform	**Min.**	**Lower Quartile**	**Median**	**Upper Quartile**	**Max.**	**Mean**	σ
Facebook	1.00	1.00	1.00	1.00	1.00	1.000	0.000
Twitter	1.00	1.00	1.00	1.00	13.00	1.062	0.637
Reddit	1.00	1.00	1.00	1.00	4.00	1.182	0.480

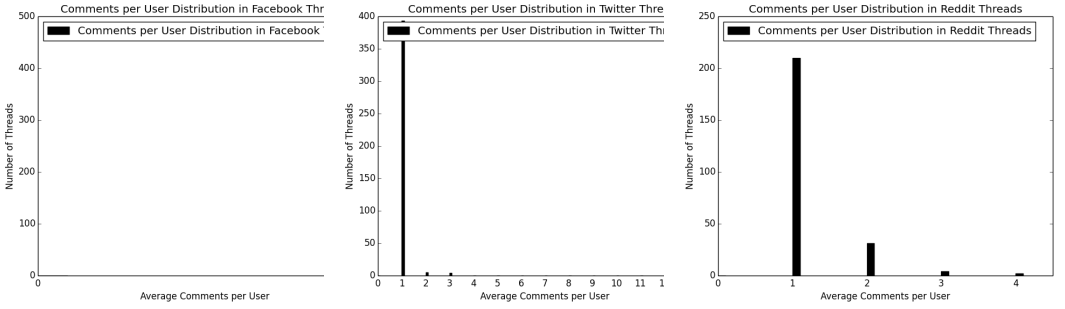


Figure 5.3: Comment per user distribution on Facebook, Twitter, and Reddit

a 'long tail' of outliers. The distribution of comment lengths for Twitter appears as a more traditional bell-curve. This is likely due to the upper-limit on comment length, restricting the total size of comments that users can post.

Table 5.3 (and Figure 5.3) shows the distribution of the total number of comments made by each user, in each thread. All three platforms had a very low number of different comments per user, Facebook in particular never exceeding 1. Twitter and Reddit both had mean and median values close to 1, with $\sigma < 0.7$ and $\sigma < 0.5$ respectively.

Table 5.4 (and Figure 5.4) shows the distribution of the total number of direct replies to other comments within each thread. Interestingly, Facebook appeared to have zero direct intra-thread

Table 5.4: Distribution of the number of internal replies in each thread

Platform	**Min.**	**Lower Quartile**	**Median**	**Upper Quartile**	**Max.**	**Mean**	σ
Facebook	0.00	0.00	0.00	0.00	0.00	0.000	0.000
Twitter	1.00	4.00	7.00	14.00	192.00	10.442	12.627
Reddit	0.00	1.00	5.00	49.00	4372.00	129.526	443.849

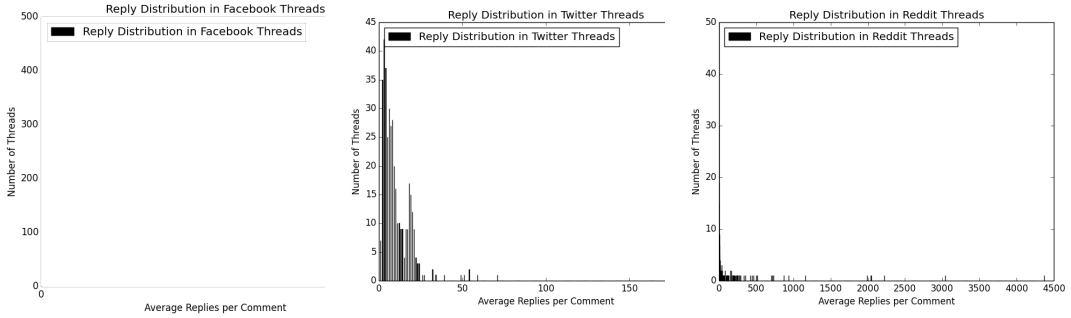


Figure 5.4: Total replies distribution on Facebook, Twitter, and Reddit

replies. Whether this is due to people unwilling to use the (at the time) new feature, or due to already having other methods (such as tagging) to denote replies is not known.

Threads that fell outside of the interquartile range of any of these categories were removed from the sample to remove outlying data points. This left 139 Facebook threads (with a total of 14,556 individual posts), 113 Twitter threads (with a total of 1,021 individual posts), and 52 Reddit threads (with a total of 1,123 individual posts).

Each of these "normal" threads was then annotated by topic, using one or more of the following seven categories based on an aggregation of the topics provided by the five news distribution sites the social media accounts of which the stories were sourced from: *Current Events*, stories covering recent or ongoing occurrences; *Business*, stories involving businesses and/or the economy; *Politics*, stories relating to politics, politicians, elections, etc.; *Science & Technology*, stories relating to science, the environment, new technologies, etc.; *Entertainment & Arts*, stories involving cultural events, celebrities, etc.; *Sports*, stories covering sporting events and participants; and *Features*, stories that form an opinion piece, or focus on an individual rather than the broader story they are a part of. The distribution of these topics is shown in Table 5.5.

To reduce the variable dimensions of the examined posts, and the complexity of the annotation task, the three most common topics (that were represented by at least one thread on each platform) – namely Current Events, Politics, and Entertainment & Arts – were selected, and (for each platform) one thread in each topic was selected at random until the total number of posts on that topic was above 400, or there were no more threads to select. Statistics of the final sample, including number of posts and users, are shown in Table **??**.

Table 5.5: Threads by topic

Platform	Current Events	Business	Politics	Science & Technology	Entertainment & Arts	Sports	Feature
Facebook	52	18	23	16	20	3	32
Twitter	53	3	24	12	12	3	20
Reddit	41	4	2	6	2	0	0
Total	146	25	49	34	24	6	52

Table 5.6: Metrics of discussions sampled from Twitter, Facebook and Reddit

Metric	Facebook	Twitter	Reddit	Total
Posts	629	357	254	1240
Direct replies	0	110	162	272
Number of users	609	313	166	1088
Average posts per user	1.03	1.14	1.53	1.14
Average words per post	18.54	13.08	29.91	61.53
Average characters per post	104.29	82.44	185.20	371.93

5.1.2 Annotation

A classification system was devised based on the previous extensions proposed to the expert review in Section 4.5. For this, only the type of nodes determined present in each locution were required to be marked. The relationships between nodes were not taken into account for this data set, due to the extreme cost of fully annotating threads at this scale. However, this is a cost-effective trade-off, as examining individual comments is sufficient to determine whether the categorisation system is complete. The possible categories were as follows: Information, Transition, Logical Attack, Logical Support, Rhetorical Attack, Rhetorical Support, Preference, Persona, Group, Audience, Implied Relationship, and Implied Belief. As before, annotations were carried out by an individual researcher (the author).

During the annotation process, it was deemed necessary to introduce three other categories to account for circumstances not otherwise covered. These categories are as follows: Spam/Advertisement, which encompasses (often automated) posts that promote a business or service unrelated to the discussion; Unknown, which encompasses posts in a language unknown to the annotator, broken or corrupted text or emojis, or any other circumstances in which the intent of the post cannot be reasonably assumed; and None, encompassing circumstances when a post is totally blank, or blank apart from "tagging" a user (to alert them of the thread).

Table 5.7: Classifications present per platform (raw figures)

Classification	Facebook	Twitter	Reddit	Total
Information	234	242	201	677
Transition	24	26	21	71
Logical Attack	33	52	84	169
Logical Support	4	10	19	33
Rhetorical Attack	310	117	49	476
Rhetorical Support	183	73	41	297
Preference	2	1	2	5
Persona	365	116	55	536
Group	42	37	12	91
Audience	85	44	30	159
Implied Relationship	5	2	0	7
Implied Belief	20	12	6	38
Spam/Advertisement	18	6	0	24
Unknown	22	9	6	37
None	31	10	1	42

5.2 Data Analysis

Table 5.7 shows the total number of annotations per classification, broken down by platform. In total, the most commonly used annotation was, perhaps unsurprisingly, Information, followed closely by Persona. Rhetorical Attack and Support, as well as Logical Attack follow, with Logical Support being used much less, comparatively. The least used annotations were Implied Relationship and Preference (both coming behind Spam/Advertisement).

The relative proportions of annotations are shown in Table 5.8. Information nodes are present within an overwhelming majority of Reddit posts (almost 80%), and are comparatively sparse within Facebook posts (less that 40%). Reddit posts also had a comparatively higher number of Logical Attacks and Supports, and fewer Rhetorical Attacks and Supports, than either of the other platforms. This suggest that discussions on Reddit are more dialectic than on Facebook or Twitter.

Table 5.9 shows the number of annotations present alongside n other annotations. For example, posts are annotated as containing Information and nothing else in 216 cases, Information and at least one other classification in 183 cases, Information and at least two other classification in 169 cases, and so on. Most annotations appear most frequently in conjunction with other annotations. Information annotations appear frequently on their own (but also frequently together with other annotations). Persona, Group, and Audience (in this sample) never appear without at least one other classification being present, as they are almost acted upon by another form of relationship, such as an Implied Relationships or Implied Belief, which also never appear alone. This is also true of Logical Support and, in most cases, Logical Attack, as some form of Information is usually present as evidence of the claim (except in cases where the statement is a simple negation/affirmation).

Table 5.8: Classifications present per platform (as percentage of posts)

Classification	**Facebook**	**Twitter**	**Reddit**	**Total**
Information	37.20%	67.79%	79.13%	54.60%
Transition	3.82%	7.28%	8.27%	5.73%
Logical Attack	5.25%	14.57%	33.07%	13.63%
Logical Support	0.64%	2.80%	7.48%	2.66%
Rhetorical Attack	49.28%	32.77%	19.29%	38.39%
Rhetorical Support	29.09%	20.45%	16.14%	23.95%
Preference	0.32%	0.28%	0.79%	0.40%
Persona	58.03%	32.49%	21.65%	43.23%
Group	6.68%	10.36%	4.72%	7.34%
Audience	13.51%	12.32%	11.81%	12.82%
Implied Relationship	0.79%	0.56%	0.00%	0.56%
Implied Belief	3.18%	3.36%	2.36%	3.06%
Spam/Advertisement	2.86%	1.68%	0.00%	1.94%
Unknown	3.50%	2.52%	2.36%	2.98%
None	4.93%	2.80%	0.39%	3.39%

Table 5.9: Classifications present alongside exactly *n* other classifications

Classification	**Number of Other Classifications Present**							
	0	**1**	**2**	**3**	**4**	**5**	**6**	
Information	216	183	169	64	34	6	3	2
Transition	22	30	10	2	4	3	0	0
Logical Attack	2	112	18	31	3	1	1	1
Logical Support	0	24	4	3	0	1	0	1
Rhetorical Attack	48	172	142	65	36	8	3	2
Rhetorical Support	30	74	33	84	62	9	3	2
Preference	0	2	2	1	0	0	0	0
Persona	0	196	133	131	62	9	3	2
Group	0	15	34	26	10	2	2	2
Audience	0	0	11	80	56	7	3	2
Implied Relationship	0	0	2	5	0	0	0	0
Implied Belief	0	1	15	17	4	1	0	0
Spam/Advertisement	23	1	0	0	0	0	0	0
Unknown	37	0	0	0	0	0	0	0
None	42	0	0	0	0	0	0	0

Table 5.10 goes into further depth, showing a two dimensional matrix of the frequency with which particular pairs of annotations appear in the sample. When comparing this with the data from Table 5.9, we can see that Information annotations are most frequently paired with Rhetorical Attacks and Personas (due to Rhetorical Attacks acting on Personas). Logical Attack and Logical Support annotations are most frequently paired with Information, presenting evidence of claims being a core aspect of a logical tactic. Rhetorical Attack and Rhetorical Support annotations are most frequently paired with Persona, Group and Audience; again, these more social tactics are almost by definition more heavily linked with the individuals and groups that make up the discussion.

Table 5.10: Frequency of tactics occurring together in Posts annotated as having (at least) two classifications

	Information	**Transition**	**Logical Attack**	**Logical Support**	**Rhetorical Attack**	**Rhetorical Support**	**Preference**	**Persona**
Information	N/A	25	159	31	217	90	4	215
Transition	25	N/A	9	0	15	7	0	14
Logical Attack	159	9	N/A	4	43	7	1	34
Logical Support	31	0	4	N/A	5	3	0	6
Rhetorical Attack	217	15	43	5	N/A	61	2	330
Rhetorical Support	90	7	7	3	61	N/A	0	249
Preference	4	0	1	0	2	0	N/A	1
Persona	215	14	34	6	330	249	1	N/A
Group	66	1	9	2	80	19	1	29
Audience	42	8	2	1	52	138	0	145
Implied Relationship	5	0	0	0	3	2	0	4
Implied Belief	22	2	1	0	22	4	0	16
Spam/Advertisement	1	0	0	0	0	0	0	0
Unknown	0	0	0	0	0	0	0	0
None	0	0	0	0	0	0	0	0

	Group	**Audience**	**Implied Relationship**	**Implied Belief**	**Spam/Advertisement**	**Unknown**	**None**
Information	66	42	5	22	1	0	0
Transition	1	8	0	2	0	0	0
Logical Attack	9	2	0	1	0	0	0
Logical Support	2	1	0	0	0	0	0
Rhetorical Attack	80	52	3	22	0	0	0
Rhetorical Support	19	138	2	4	0	0	0
Preference	1	0	0	0	0	0	0
Persona	29	145	4	16	0	0	0
Group	N/A	7	4	14	0	0	0
Audience	7	N/A	0	21	0	0	0
Implied Relationship	4	0	N/A	1	0	0	0
Implied Belief	14	21	1	N/A	0	0	0
Spam/Advertisement	0	0	0	0	N/A	0	0
Unknown	0	0	0	0	0	N/A	0
None	0	0	0	0	0	0	N/A

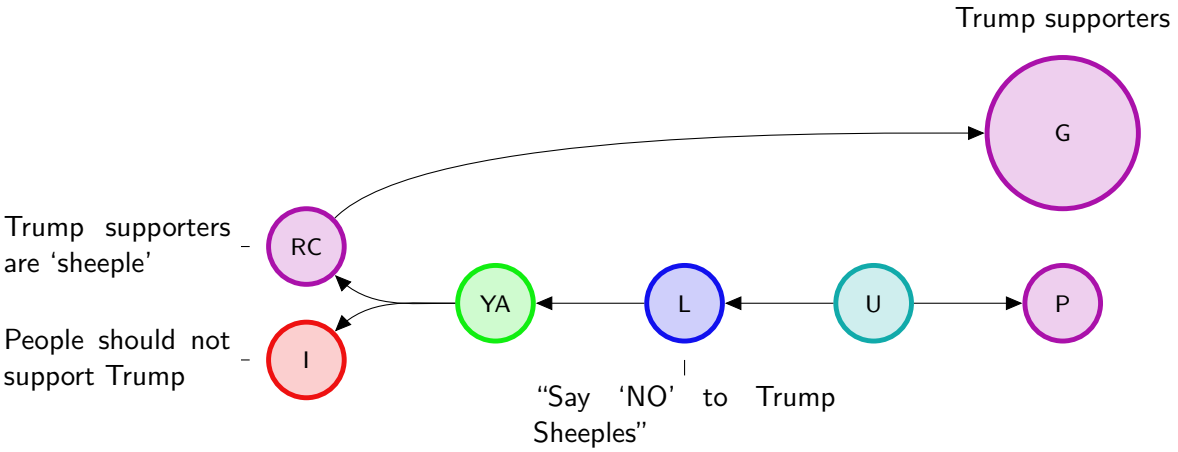


Figure 5.5: Post from Facebook case study, showing rhetorical attack against a group

5.3 Narrative Account

In this section, one thread was randomly chosen from each of the social media platforms to be examined as a case study. These three threads are reproduced in full in Appendix D.

5.3.1 Facebook

Appendix D.1 shows a thread on Facebook, starting with the post:

> *"US election 2016: Trump wins Nevada - Strengthening his position in the Republican presidential race."*

The overall structure of this thread is very flat, with no nested replies (though, through observation, participants do still refer to previous posts, or "tag" other users by referring to them by name). The majority of users post a comment sharing their own views, predominantly supporting or attacking Trump. Often, these posts are self-contained, relating to no other information in the thread aside from the original article. These range from relatively short, such as:

> *"Say 'NO' to Trump Sheeples"* (Figure 5.5)

or:

> *"Yes trump love it .from the UK"* (Figure 5.6)

which simply voice support or opposition to Trump as a candidate, and implicitly signal the user's own political beliefs. The latter example highlights that a Persona does not necessarily

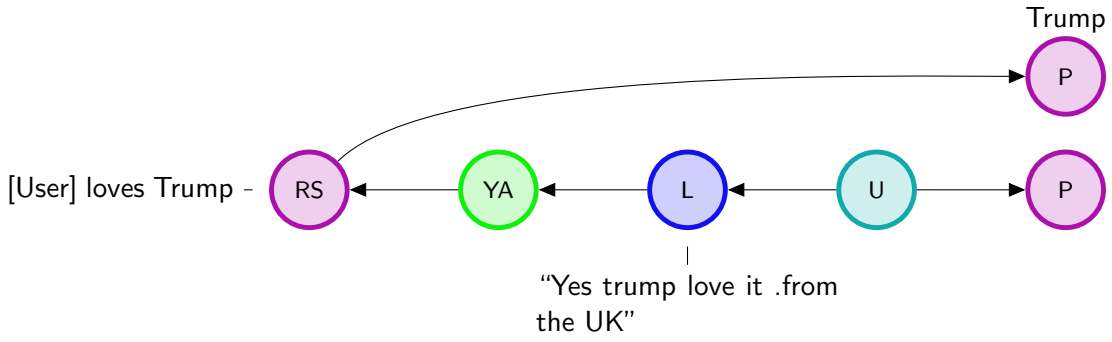


Figure 5.6: Post from Facebook case study, showing rhetorical support for a persona

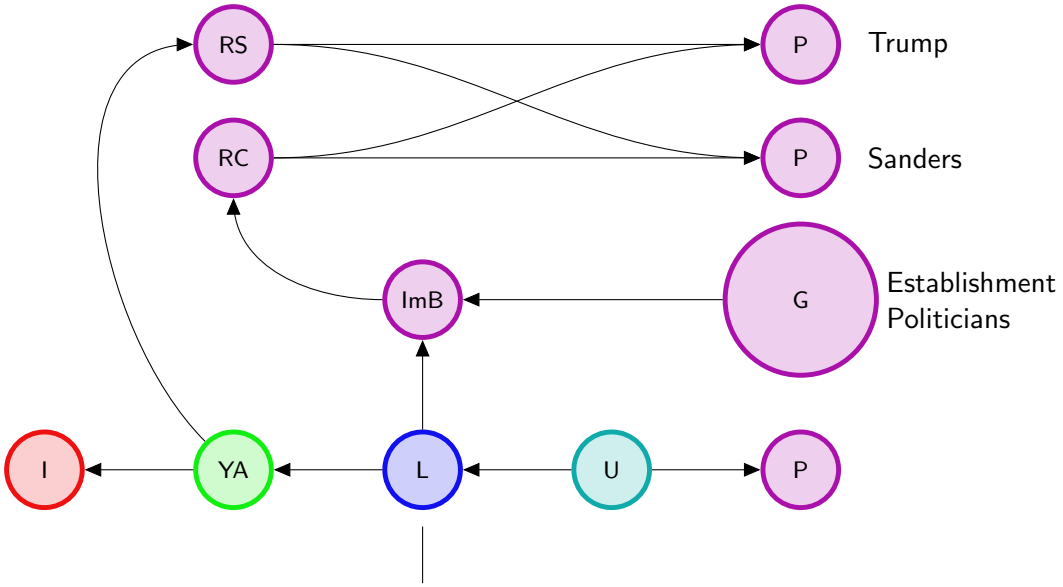
need to be attached to an active UserAccount to be attacked or supported. This leads to the interesting theoretical possibility of wholly fabricated "strawman" Personas being constructed by members of the discussion.

In contrast, some users post detailed and articulate arguments, explicitly spelling out why they believe a certain way. here, the participant uses the social force of an undesirable group (in this case, "establishment politicians") to justify their support for the two anti-establishment candidates:

> *"Interesting times we live in. I'd really like to see Trump and Sanders win their respective party's nomination. I think they might both force our two political parties to reform and produce better candidates. It'd be nice to see them shake up DC too, they're both political outsiders and it seems most establishment politicians aren't very fond of either of them."* (Figure 5.7)

A number of users post comments simply to tag other users (presumably friends) to notify them of the story. Some users, possibly bots, also use the thread to advertise t-shirts, or other seemingly unrelated news-stories, likely due to the large user-density of Facebook. Broadly speaking, due to the lack of explicit intra-thread replies, it appears that users mostly use Facebook to signal their own views, without directly engaging in discussion with other users in the same thread, while at the same time using it to highlight particular posts of interest to friends who may not otherwise see them.

Figure 5.8 plots the increase of different tactics over time, and shows that the vast majority of comments made are rhetorical in nature, and the vast majority of those are attacks. However, these attacks are rarely aimed at other participants in the discussion; they are predominantly aimed at the subject of the initial post, and the topic at hand; topics about a particular individual appear to draw a large amount of ire from the social web, in particular one as divisive as a politician. Figure 5.9 shows the breakdown of contributions by user, which conveys that most users make some form of rhetorical comment; it also highlights that throughout the entire thread, few users posts more than once.



"Interesting times we live in. I'd really like to see Trump and Sanders win their respective party's nomination. I think they might both force our two political parties to reform and produce better candidates. It'd be nice to see them shake up DC too, they're both political outsiders and it seems most establishment politicians aren't very fond of either of them."

Figure 5.7: Post from Facebook case study, showing a more complex argument, drawing on a number of ASWO features

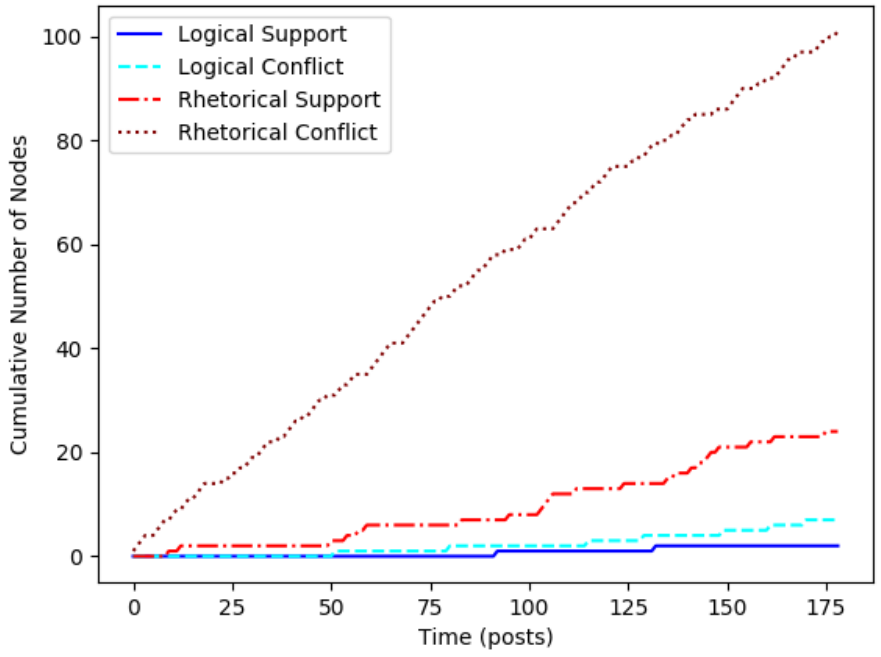


Figure 5.8: Cumulative logical and rhetorical tactics over time on Facebook

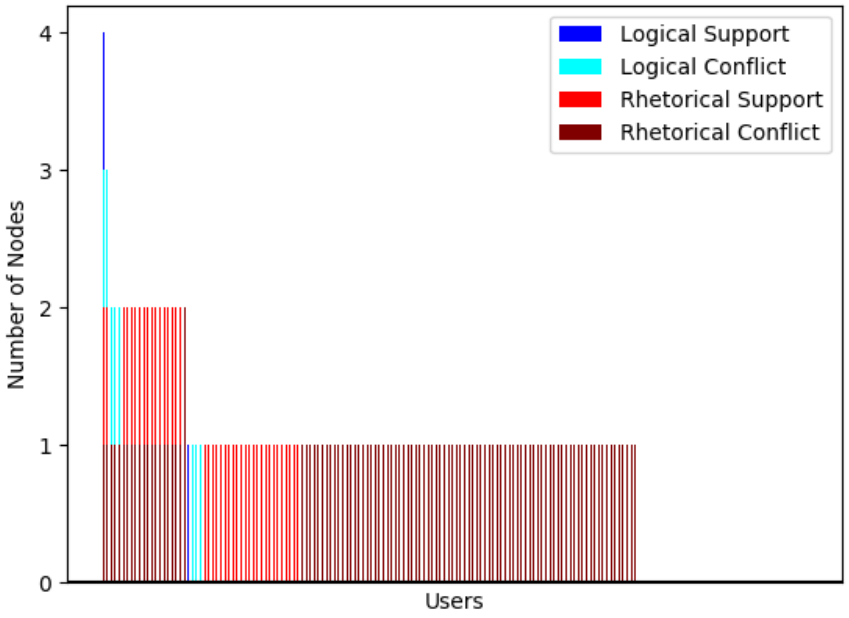


Figure 5.9: Logical and rhetorical tactics per user on Facebook

5.3.2 Twitter

Appendix D.2 shows a thread on Twitter, starting with the post:

> *"Kesha 'in tears' after judge denies her release from Sony where producer allegedly raped her* `https://t.co/thYQgcH9AW https://t.co/XBnz5x5fIJ`*"*

Here, more intra-thread replies can be observed, with users directly and explicitly replying to one-another, either to make a counterpoint (using an *argumentum ad baculum* threat):

> *"@Btwsts @DailyMailCeleb you do realise thats slander without proof, If sony wanted too they could sue you. People have been sued for less."* (Figure 5.10)

or simply to request more information on the topic:

> *"@Btwsts what rape?"*

(if taken literally; read another way, this post could instead be voicing disbelief/scepticism, again highlighting the difficulties of modelling when faced with a lack of tone or other familiar vocal or physical language cues).

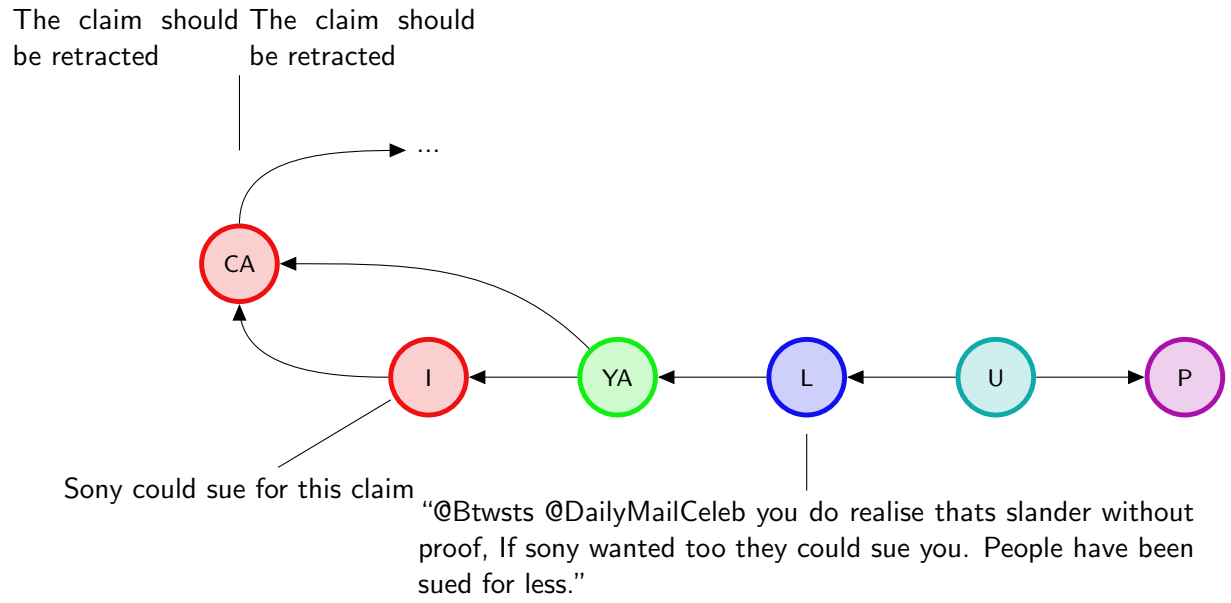


Figure 5.10: Post from Twitter case study, showing a threatening *argumentum ad baculum*

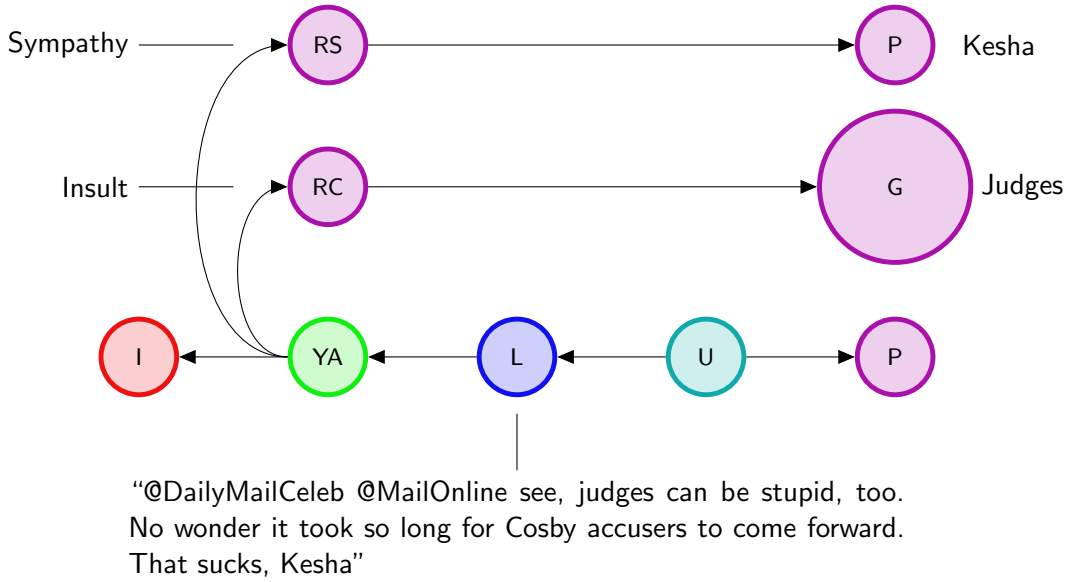


Figure 5.11: Post from Twitter case study, showing combined (sympathetic) rhetorical support, and (insulting) rhetorical attack

Users also take the opportunity to express their sympathy for Kesha, and their displeasure at the judges.

> *"@DailyMailCeleb @MailOnline see, judges can be stupid, too. No wonder it took so long for Cosby accusers to come forward. That sucks, Kesha"* (Figure 5.11)

One user also reposts the entirety of the original post, verbatim:

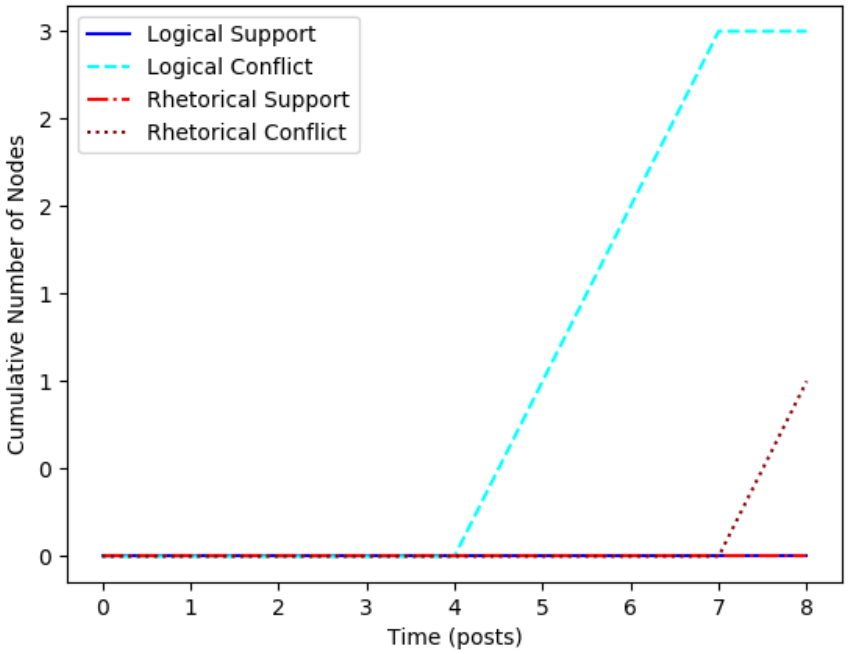


Figure 5.12: Cumulative logical and rhetorical tactics over time on Twitter

> *"RT @DailyMailCeleb: Kesha 'in tears' after judge denies her release from Sony where producer allegedly raped her* `https://t.co/thYQgcH9AW` *ht"*

By doing this, they shared this content with their own followers as well. However, by quoting and reposting in this way (rather than just "retweeting", which copies the original tweet to their followers) they also explicitly chose to reply to the original post, which is not required.

Figures 5.12 and 5.13 highlight the lack of intra-thread engagement in this particular thread. While this is partly due to it being the shortest of the three observed case studies, there are still only three users present that engage in any sort of support or attack (logical or rhetorical); the other users either provide information or request it.

5.3.3 Reddit

Appendix D.3 shows a thread on Reddit, starting with the post:

> *"Brexit against Scotland's wishes would 'almost certainly' trigger independence referendum, warns Nicola Sturgeon"*

This thread had many more internal replies than either of the others, with discussion moving back and forth between different "sub-threads": Reddit, unlike the two previously observed platforms highlights replies by indenting them within the body of the page for greater clarity.

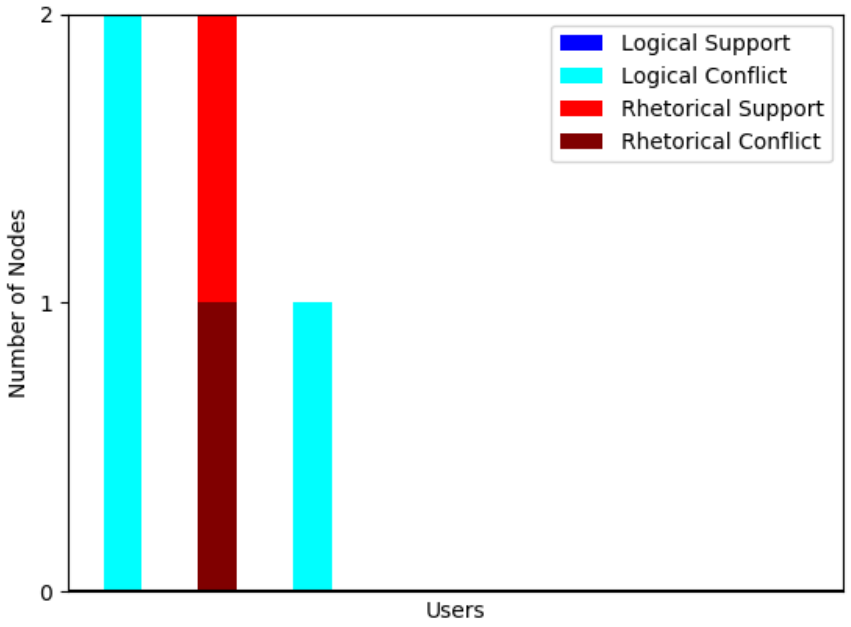


Figure 5.13: Logical and rhetorical tactics per user on Twitter

Comments routinely refer directly to the comments they reply to, for example:

> *"You actually can call for a referendum whenever you want."*

(Figure 5.14)

> *"Like France, Ireland etc."*

(Figure 5.14)

or in other instances:

> *"that arsey dwarf is (one of) the reasons why we've got a tory government"* (Figure 5.15)

> *"If every single person in Scotland voted labour, the Tories still would have won"* (Figure 5.15)

> *"the snp claimed labour would need them to govern, and that scotland would have the sway over a future labour government.*

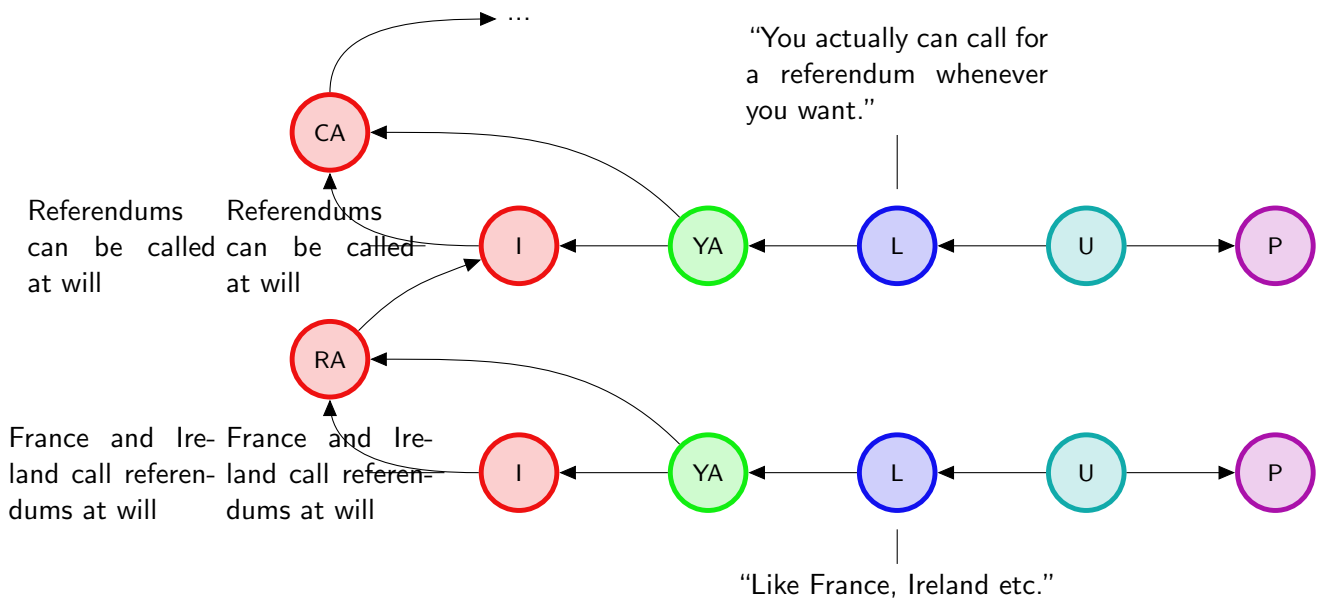


Figure 5.14: Post from Reddit case study, showing the logical support of a previous argument through example

> *great for her home audience, but the tories used it to make labour look like their puppet- a british parliament working for scotland.*
>
> *so scotland gets all those mps but as you say, not enough to decide any policies AND they ensured a tory victory.*
>
> *edited for my shocking spelling"* (Figure 5.15)

This highlights that, on Reddit, people are more inclined to argue or discuss in good-faith, even when resorting to crude humour or vulgar language. This is perhaps unsurprising: as users must specifically browse to this subreddit, rather than come across such posts by chance.

Interestingly, this thread also contains a post from a (commonly used, and self-acknowledged) bot, that automatically provides a summary of the article reference in the original post, for users too busy or unwilling to read the entire article, allowing them to still make a more informed contribution to the discussion than they might otherwise be able to.

Figure 5.16 shows that on this platform too there is still a clear prevalence of attack over support, as might be expected in an argument. However, in this topic at least, logical attacks are favoured over rhetorical attacks and, within rhetorical tactics, support is favoured over attack. Figure 5.17 shows that some users choose to post only a single comment, using various tactics, others engage further, posting multiple times and using an array of different tactics, both logical and rhetorical.

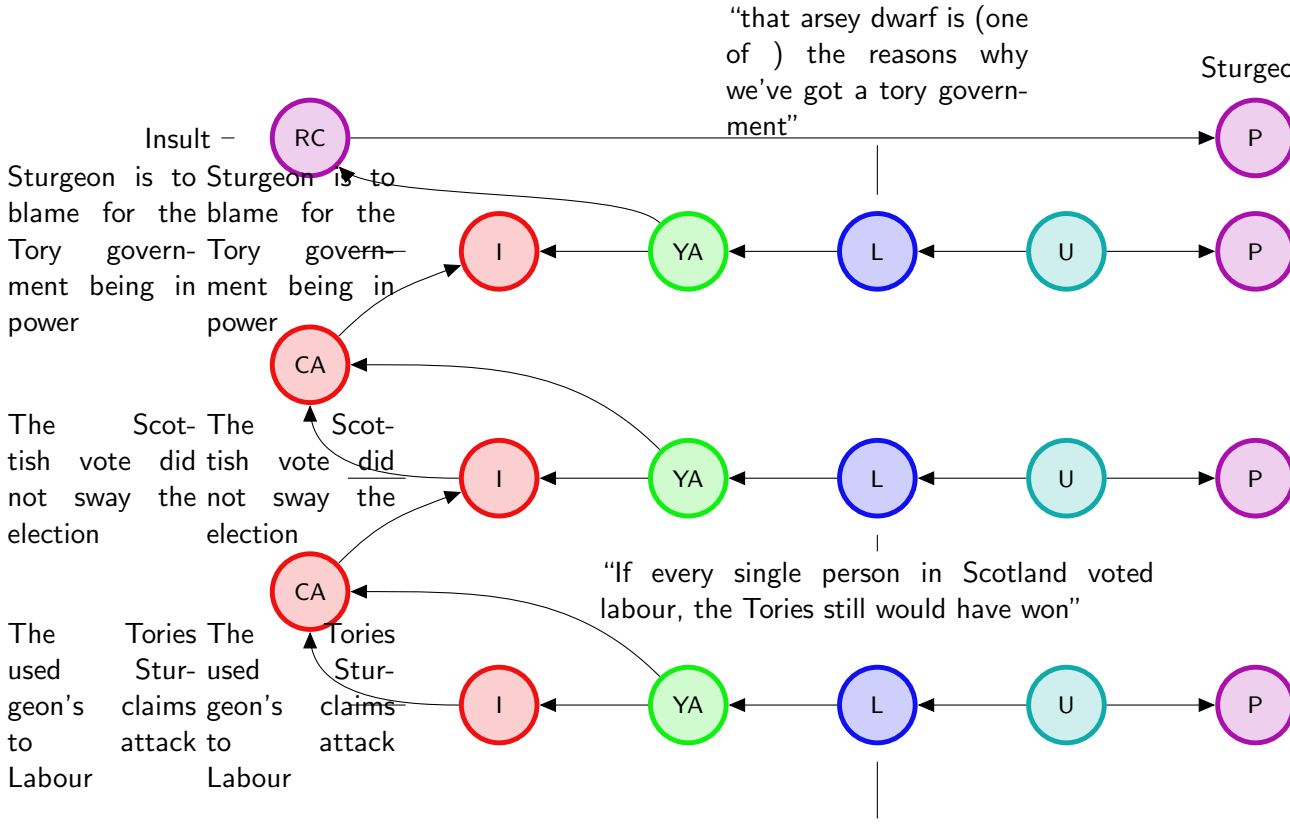


Figure 5.15: Post from Reddit case study, showing a more complex argument chain

5.4 Summary

This chapter examined the structure of typical social media threads across three different platforms by collecting a dataset of 1,500 threads that were then pruned to ensure they were discursive, and to remove outliers. This left 304 individual threads containing 16,700 posts, 1,240 of which were then annotated with a classification system based on the ontological framework developed in Chapter 4. These threads were compared them with one another, both in terms of structure (i.e., length of comments, number of replies, etc.) and in terms of annotations present. A set of individual case-studies were also examined, highlighting the observable differences in

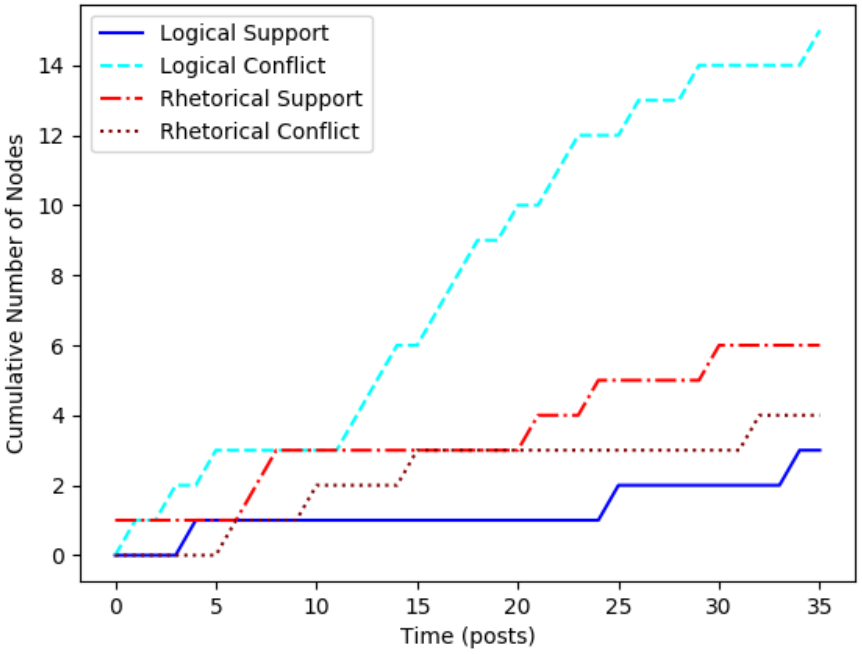


Figure 5.16: Cumulative logical and rhetorical tactics over time on Reddit

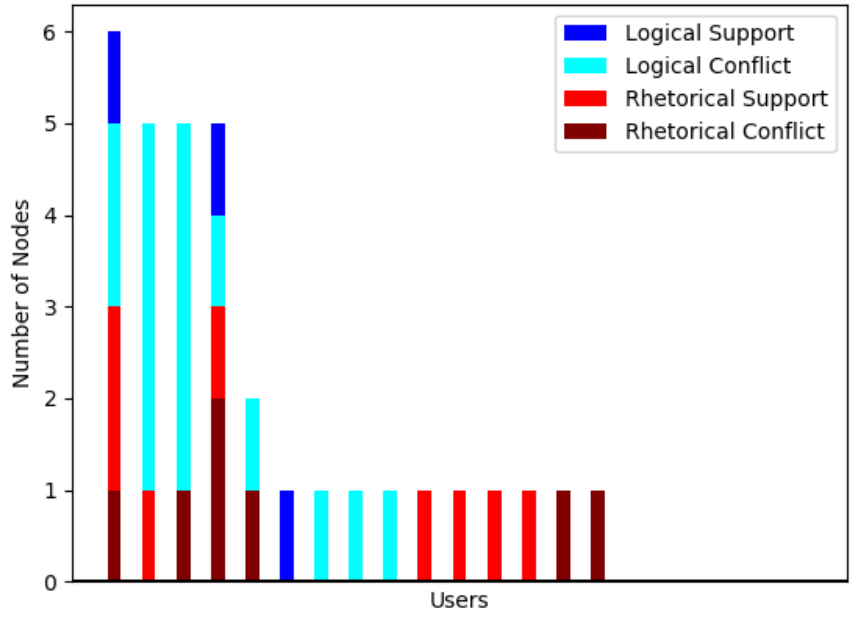


Figure 5.17: Logical and rhetorical tactics per user on Reddit

types of argumentation between platforms. These individual case studies in particular highlight the ability of the ASWO to adequately model social eristic rhetorical argumentation (*R*3).

The case studies illustrate the differences that can be seen between discussions carried out on different social media platforms; however, due to the limitations put on topics considered, and the sheer magnitude of discussions that are carried out across the social web, these case studies cannot be used to justify broad statements about every facet of social discussion. With this in mind, Rhetorical Support and Attack appeared to be much more prevalent than any Logical techniques, particularly posts on Facebook, which frequently made some form of attack at the subject of the initial post (or the topic in general), and then leaving without further discussion: a "fire and forget" approach. The exception to this was Reddit, which displayed a higher proportion of Logical Attacks than any other tactic, often seemingly more proactively engaging with other users in discussion, rather than (solely) the initial topic. This is likely due to the fact that on Facebook or Twitter a user is presented with a stream of subscribed or recommended content, Reddit users must deliberately seek out these topics and are therefore more likely to engage with them and other such users.

The implications of these finding appear to confirm that communities that seek out discussion are more likely to engage with them in good faith, whereas those that a user is obligated to see through curated or recommended "feeds" (even if not actually forced to engage with), is more likely to elicit a more rhetorical response. Some of this, obviously, will be down to the individual culture of the platform (and there will of course be exceptions to this). The impact of a user seeing a vast number of offensive (or, conversely, entertaining) posts (even if they do not directly engage with other participants) will almost certainly shape the way that they make any subsequent contribution, and thus it is important for annotators to bare this in mind when modelling, and imperative that they should have the tools to model it.

To further support modellers of eristic argumentation, it is important to provide an understanding of how participants react to different types of argumentation. To achieve this, the following chapter addresses the final research question of this thesis (*R*4), in which the annotated set of posts used in this chapter forms the basis of an experiment into how different types of argument are perceived by their audience, and the value that participants attribute to logical and rhetorical statements within discussions on the social web.

Chapter 6

Perception of Rhetorical Tactics in Individual Comments

Having developed a framework for categorising social media post as including different types of argumentative tactics in Chapter 5 based on the ASWO, this chapter seeks to answer the final research question *R*4 (*Do rhetorical techniques affect the way in which users perceive and engage with online argument?*); although the ASWO is capable of modelling both logically and rhetorical argumentation, because the types of argumentation it is designed to aid in modelling are extremely difficult to automatically reason over, it is important for both annotators and analysts to understand the way in which people exposed to different kinds of argument think about and respond to them, so that they may better understand the way in which arguments develop on the web.

Perception can be thought of as the way in which users understand the tone, persuasiveness, entertainment value or information content of an argument (Sundar, 2000). Engagement, on the other hand, can be thought of as how likely users are to react to, respond to, or participate in the argument itself, and they ways in which they do so. This is not limited to replying to a post: users of social media can engage in multiple ways, including replying, sharing or voting (Markova and Petkovska-Mirčevska, 2013).

To examine how different categories of argumentation may affect a user's perception (and subsequent engagement), an experiment was carried out in which participants were shown a number of social media posts from the dataset acquired in Chapter 5, and asked a series of questions correlating to perception and engagement (answerable on a standard Likert scale) about each of them. These questions gauged both how they felt about the post (e.g. persuaded, entertained, or offended), as well as how they would act on it given the opportunity (e.g. replying, sharing, or reporting).

6.1 Experimental Hypothesis

The working experimental hypothesis for this study was that participants would perceive (and subsequently report their likely engagement with) social media posts containing logical tactics in a significantly different manner than those containing rhetorical tactics.

6.2 Methodology

6.2.1 Data Sample

This work used the annotated "normal" threads, the acquisition of which is described in Section 5.1.1.

6.2.2 Participants and Survey

Ethical approval for this study was granted under application number 18149; the application (including participant information forms) is detailed in Appendix A.2.

Participants were invited via a general call on social media; they were provided an information sheet and consent form, and asked to confirm that they were above the age of eighteen. Participants were each shown a selection of eighteen posts taken from the sample described in Section 6.2.1 (three different news-types across three social media platforms). For context, they were also show the initial post in the thread and the post that had the target post as a direct reply (if any).

Participants were asked to rate each comment they were shown on a Likert-scale of nine different questions, covering aspects of both argumentation and perception, and social media and engagement.

The basis for the questions pertaining to argumentation and perception are determined as follows. Firstly, and perhaps obviously, a post must be legible and comprehensible enough enough to be understood, to make a useful contribution to the discussion; therefore participants were asked to rate the overall coherence of the comment. Secondly, as the presence of data and claims makes up a large part of argumentation modelling (Toulmin, 1958; Chesñevar et al., 2006; Reed et al., 2008) and the need in particular, for a compelling argument, for claims to be considered to be believable (Wigmore, 1913; Metzger and Flanagin, 2013), participants were asked to rate the comment on its (credible) information content. Common theories of argument focus on persuasion of the audience as the overall goal and intended outcome of the discussion (Payne and Vaughan, 1969; Schopenhauer, 1831; Schneider et al., 2014); as such, participants were asked how persuasive they felt the comment was. To account for the notion that eristic argumentation is commonly undertaken as a recreational pursuit, or for the purpose of catharsis (Schneider

et al., 2014) participants were also asking how entertaining they found each comment. Finally (in terms of argumentation and perception), was the notion of incivility and offence. Generally, incivility in online discussions can polarise the opinion and perception of the audience (Anderson et al., 2014); this is addressed with a question to determine whether participants considered the comment to be offensive.

The basis for the questions pertaining to social media and engagement is taken from the work by Kietzmann et al. (2011), who describe seven different dimensions of interaction and engagement on the social web: *conversations*, *sharing*, *presence*, *reputation*, *identity*, *relationships*, and *groups*. *Conversation* describes the extent with which users converse between themselves; this obviously forms a fundamental part (perhaps *the* fundamental part) in maintaining the discourse. If no users were to respond, there would be no discourse at all, and users would likely move on to other threads, leaving the post to be quickly forgotten on the more ephemeral platforms, and those that use algorithms to curate the posts that their users see (Langlois and Elmer, 2013). To judge this element of social engagement, participants were asked how likely they would be to reply to the comment with one of their own. *Sharing* describes the extent to which users exchange content, and is also a key element of many social media platforms, particularly with regard to information or knowledge sharing and collaboration (Hermida et al., 2012; Panahi et al., 2012). This dimension of social engagement is judged by asking participants how likely they would be to share the comment with their friends or followers. *Presence* describes a user's digital footprint; whether they are, for example online or offline, their role, and how easily (or whether) their content can be seen by other users. This aspect is evaluated in the study by a question asking participants if they would report the presented comment to the platform's moderators, thus hiding it from their view (and potentially the view of all other users if the report resulted in a ban for the author). *Reputation* describes a users social standing in the community, which can be implicit (determined from evaluating the content they publish, context-cues that may imply the user is a bot, etc.) or explicit (determined through a metric such as upvotes or downvotes). An explicit reputation score can influence the trust in a particular user (and the content which they post) (Kietzmann et al., 2012), and may also have an impact on how many other people see the post (Bucher, 2012). This feature of social engagement is evaluated in this study by asking participants if they would up- or down-vote the presented comment. *Identity* describes the way in which users present themselves (through their profile or avatar, etc.) and is disregarded for this study, as the focus is solely on how the arguments themselves are perceived, rather than the contributors or the audience. However, this is a topic that would be interesting to study more deeply, particular with regard to how the identity of the user posting the argument affects the perception of the argument itself (Blount et al., 2015b). Finally, the topics of *relationships* and *groups* describe how users form social bonds with other users, whether individually or in communities. These are also disregarded for the purposes of this study; while these would also be of interest for further study (to determine what connections are made as a result of the discussion), this study is only concerned with direct engagement to the discussion itself.

The full set of questions therefore (with an associated shorthand) were as follows:

1.	This comment is coherent/easy to understand	*Co*
2.	This comment contains (or appears to contain) credible information	*Cr*
3.	This comment makes (or attempts to make) a persuasive argument	*P*
4.	This comment is (or attempts to be) entertaining	*E*
5.	This comment is (or attempts to be) offensive	*O*
6.	Would you be more or less likely to reply to this comment than average?	*Rpl*
7.	Would you be more or less likely to share this comment (to friends/followers/etc.) than average?	*S*
8.	Would you be more or less likely to up-/down-vote this comment than average?	*V*
9.	Would you be more or less likely to report this comment than average?	*Rpr*

Participants were instructed to answer the first five questions on a Likert-scale of *Strongly Disagree*, *Disagree*, *Neutral*, *Agree*, and *Strongly Agree*. The latter four questions were answered on a scale of *Very Unlikely*, *Unlikely*, *Neutral*, *Likely*, and *Very Likely*. Participants were also given an optional free-text field in which they could justify their answers if they chose. Participants were instructed to skip any comments that they felt they could not accurately rate. These questions were presented to participants via a web-survey, a screenshot of which is shown in Figure 6.1. Participants were also cautioned against following any URLs present in the comments.

In addition to the eighteen randomly selected posts, every participant was also shown (and asked to rate) two additional posts (common to all participants), to judge the overall inter-rater agreement. These 'control' posts were selected at random from the sampled pool of comments and manually inspected to ensure they were non-empty and comprehensible (e.g., English-language).

In all, 60 participants responded to the survey, with 33 completing it fully and 27 partially completing it. These participants were shown 436 unique comments, with 746 questions answered (giving 6714 total ratings).

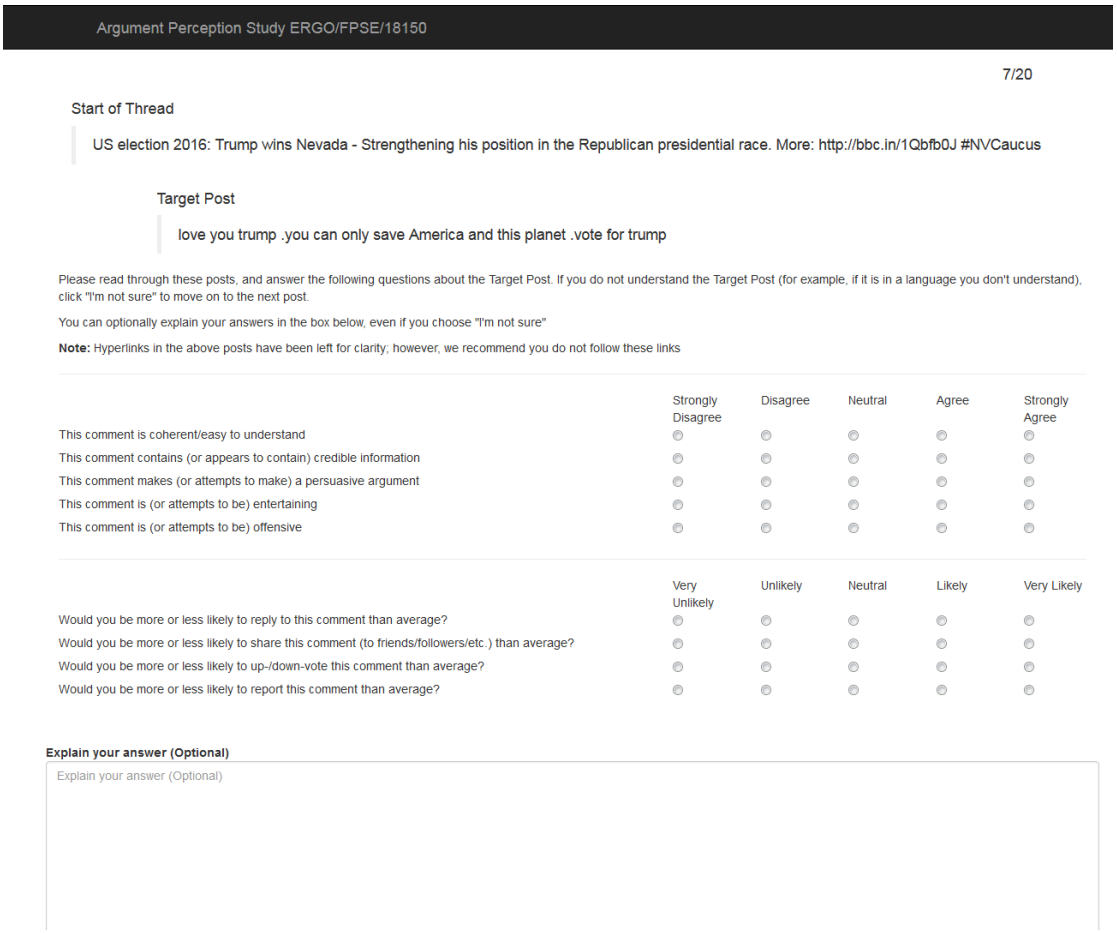


Figure 6.1: The questionnaire as presented to participants

6.3 Data Analysis and Results

6.3.1 Raw Data

Table 6.1 shows the raw answers that participants gave to the questions overall. There are a number of observations that can be made based on this data, when looking at the average ratings of posts as a whole.

Firstly, there is clear bimodal distribution of responses relating to the coherence of posts, suggesting that comments were either relatively easily determined to be either clear enough to read and understand (making up the majority of posts that were rated), or too opaque to make sense of.

Secondly, participants tended to find that while, on average, comments were not a credible source of information, they still attempted to make a persuasive argument to their audience.

Another bimodal distribution can be observed related to entertainment: comments were found to be either relatively entertaining, or relatively lacking in entertainment. This again suggesting that participants had a clear idea of what they deemed entertaining.

Table 6.1: Breakdown of answers given for each question

	Score				
Question	**1**	**2**	**3**	**4**	**5**
This comment is coherent/easy to understand	48	134	71	374	119
This comment contains (or appears to contain) credible information	127	230	224	133	32
This comment makes (or attempts to make) a persuasive argument	111	172	159	252	52
This comment is (or attempts to be) entertaining	126	189	151	220	60
This comment is (or attempts to be) offensive	165	208	194	143	36
Would you be more or less likely to reply to this comment than average?	232	216	183	108	7
Would you be more or less likely to share this comment (to friends/followers/etc.) than average?	291	226	176	51	2
Would you be more or less likely to up-/down-vote this comment than average?	207	179	195	143	22
Would you be more or less likely to report this comment than average?	296	204	211	30	5

The distribution towards offence is slightly towards the lower end of the responses. Participants, overall, felt that the majority of comments would not be considered offensive. However, the distribution is centred enough to suggest that some comments were harder to classify, likely due to the subjective nature of offence.

Lastly, the overall distribution of responses to questions regarding engagement (replying, sharing, etc.) are highly skewed towards the lower end of the responses, suggesting that the majority of participants would be disinclined to interact or engage with any of the comments that were presented here in any way whatsoever.

Table 6.2 shows the number of responses given per each different annotation type. Note that as a post may have multiple annotations, there may be (and is likely to be) overlap in the number of responses. From this we can see the distribution of annotations, combined with random selection and the ability of participants to skip questions, resulted in only one response for the *Spam/Advertisement* category, and only two for the *Preference* category. However, posts annotated as containing Logical Attack/Support and Rhetorical Attack/Support, the key element of this perception study, received over one hundred responses each.

Table 6.2: Number of responses given per annotation

Annotation	**Number of Responses**
Information	491
Transition	40
Logical Attack	146
Logical Support	27
Rhetorical Attack	256
Rhetorical Support	181
Preference	2
Persona	307
Group	55
Audience	104
Implied Relationship	7
Implied Belief	30
Spam/Advertisement	1
Unknown	4
None	10

6.3.2 Inter-Rater Reliability

To gain an additional insight of how much agreement there was between participants, and if there was particularly strong areas of agreement or disagreement, the first and last comment shown were identical to all participants. These comments were selected at random from the sample pool, but were manually inspected to ensure they were non-empty and comprehensible (e.g. English-language).

The first post, taken from Reddit, reads as follows:

> *The FSA is also irrelevant. Nusra and/or ISIS would be be the rulers if Assad collapsed. And also, NO the "FSA" is mostly islamists, most of the "secular rebels" have switched around and have become part of the SDF which is part of the YPG.*

This was annotated as having Information, and a Logical Attack.

Table 6.3 shows the raw answers that participants gave to the first group of questions when viewing the fist control comment. In addition to these results, 10 people skipped this comment. Table 6.4 shows the distribution of these answers: while there is a relatively large range between the minimum and maximum responses the standard deviation is not excessive. The greatest agreement between participants was for the statements exploring the perception of entertainment ($\sigma = 0.764$), which they found did not attempt to be entertaining ($\bar{x} = 1.673$), and credibility ($\sigma = 0.899$), finding it to be somewhat uncredible ($\bar{x} = 2.745$). The lowest agreement between participants was for the statement *This comment makes (or attempts to make) a persuasive argument* ($\sigma = 1.218$), suggesting that some participants felt it was trying to make an explicit argument, while others disagreed (perhaps feeling that it was providing a neutral view).

The second control post (and the last post that participants who completed the full set of questions were shown), was taken from Twitter and reads as follows:

> *@steve_walke23 @guardian @lisaocarroll Some beings are inhuman - ISIS atrocities don't bother you? Get off your priggish high horse.*

This was annotated as having Information and a Rhetorical Attack, directed against a Persona (in this case likely representing one of *@steve_walke23*, *@lisaocarroll*, or potentially both users at once).

Table 6.5 shows the raw answers that participants gave to the last comment they were shown. In addition to these results, 2 people skipped this question. Because fewer people viewed this post overall (dropping out of the study before they reached the final set of questions), the total number of responses to the questions for this control post is lower than that of the first control post. Table 6.6 shows the distribution of these answers. The standard deviations did not differ greatly

Table 6.3: Breakdown of answers given for the first post participants were shown

Question	**1**	**2**	**3**	**4**	**5**
This comment is coherent/easy to understand	6	23	8	16	2
This comment contains (or appears to contain) credible information	4	18	22	10	1
This comment makes (or attempts to make) a persuasive argument	7	11	8	24	5
This comment is (or attempts to be) entertaining	26	23	4	2	0
This comment is (or attempts to be) offensive	15	21	13	6	0
Would you be more or less likely to reply to this comment than average?	19	15	11	10	0
Would you be more or less likely to share this comment (to friends/followers/etc.) than average?	23	16	12	4	0
Would you be more or less likely to up-/down-vote this comment than average?	14	17	15	9	0
Would you be more or less likely to report this comment than average?	18	14	21	2	0

from the responses to the first group of questions in scale, though they do differ by question. The largest value of deviation ($\sigma = 1.159$) was lower than the highest standard deviation of the first group of questions, and was shared by the questions examining participants reactions, specifically replying to and voting on. The lowest values of deviation ($\sigma = 0.840$ and $\sigma = 0.904$ respectively) were higher than the lowest deviations of the first group of questions, and were in response to the likelihood participants would reply to this comment (the majority were in agreement that they were unlikely to ($\bar{x} = 2.548$)) and judging whether it was considered coherent (on average, participants found it neither excessively clear or unclear ($\bar{x} = 3.613$)).

Overall, there was a surprising range of disagreement between participants in terms of the minimum and maximum Likert responses selected, even within the two control posts; the range of answers for both posts often spanned the entire (or close to the entire) range of possible response values. These values highlight the inherent subjectivity involved in this task, with (for example), participants having different views on not only entertainment and offence, but also coherence. However, despite some outliers, most participants were in somewhat closer agreement, albeit still with variation between their responses ($\sigma < 1.3$). While this variation did reduce between the two control posts (across almost all questions asked), it did not close dramatically, suggesting that there is not necessarily a learning component to the task (in which people will gravitate towards similar answers to their peers), and that there will always be a degree of disagreement in a task such as this.

Table 6.4: Distribution of answers given for the first post participants were shown

Question	Min.	Lower Quartile	Median	Upper Quartile	Max.	Mean	σ
This comment is coherent/easy to understand	1.00	2.00	2.00	4.00	5.00	2.727	1.103
This comment contains (or appears to contain) credible information	1.00	2.00	3.00	3.00	5.00	2.745	0.899
This comment makes (or attempts to make) a persuasive argument	1.00	2.00	4.00	4.00	5.00	3.164	1.218
This comment is (or attempts to be) entertaining	1.00	1.00	2.00	2.00	4.00	1.673	0.764
This comment is (or attempts to be) offensive	1.00	1.00	2.00	3.00	4.00	2.182	0.955
Would you be more or less likely to reply to this comment than average?	1.00	1.00	2.00	3.00	4.00	2.218	1.107
Would you be more or less likely to share this comment (to friends/followers/etc.) than average?	1.00	1.00	2.00	3.00	4.00	1.945	0.961
Would you be more or less likely to up-/down-vote this comment than average?	1.00	1.50	2.00	3.00	4.00	2.345	1.031
Would you be more or less likely to report this comment than average?	1.00	1.00	2.00	3.00	4.00	2.127	0.916

Table 6.5: Breakdown of answers given for the last post participants were shown

Question	**1**	**2**	**3**	**4**	**5**
This comment is coherent/easy to understand	0	6	3	19	3
This comment contains (or appears to contain) credible information	4	10	12	4	1
This comment makes (or attempts to make) a persuasive argument	2	3	6	18	2
This comment is (or attempts to be) entertaining	4	14	8	5	0
This comment is (or attempts to be) offensive	1	4	3	17	6
Would you be more or less likely to reply to this comment than average?	8	7	7	9	0
Would you be more or less likely to share this comment (to friends/followers/etc.) than average?	11	12	7	1	0
Would you be more or less likely to up-/down-vote this comment than average?	8	9	7	6	1
Would you be more or less likely to report this comment than average?	10	8	12	1	0

Table 6.6: Distribution of answers given for the last post participants were shown

Question	Min.	Lower Quartile	Median	Upper Quartile	Max.	Mean	σ
This comment is coherent/easy to understand	2.00	3.00	4.00	4.00	5.00	3.613	0.904
This comment contains (or appears to contain) credible information	1.00	2.00	3.00	3.00	5.00	2.613	0.973
This comment makes (or attempts to make) a persuasive argument	1.00	3.00	4.00	4.00	5.00	3.484	0.979
This comment is (or attempts to be) entertaining	1.00	2.00	2.00	3.00	4.00	2.452	0.910
This comment is (or attempts to be) offensive	1.00	3.50	4.00	4.00	5.00	3.742	1.015
Would you be more or less likely to reply to this comment than average?	1.00	1.50	3.00	4.00	4.00	2.548	1.159
Would you be more or less likely to share this comment (to friends/followers/etc.) than average?	1.00	1.00	2.00	2.50	4.00	1.935	0.840
Would you be more or less likely to up-/down-vote this comment than average?	1.00	1.50	2.00	3.00	5.00	2.452	1.159
Would you be more or less likely to report this comment than average?	1.00	1.00	2.00	3.00	4.00	2.129	0.907

6.3.3 Question Breakdown

In this section, a summary of the responses to each of the questions is presented, compared against the annotations present on the post. Table 6.7 shows the mean Likert values for each question, and Table 6.8 shows the standard deviation from this mean. The full results are presented in Appendix E (Tables E.1-E.9). Note that as a post may have multiple annotations, there may be (and is likely to be) overlap in the number of responses.

In terms of perception (Co, Cr, P, E, O), when discounting post classifications that received few responses (e.g. Preferences and Spam/Advertisement) it can be seen that there is a reasonable level of variation between participants ($\sigma > 1$), which can be attributed to the variation inherent in different comments (even those with similar classifications), as well as the subjectivity of participants responses (as seen in Section 6.3.2). While there is little that can be done to account for the inherent subjectivity of argument (what one annotator might find offensive, another annotator may find entertaining), it also suggests that the use of these classifications (and their associated forms within the ASWO) can only be applied at a relatively broad level. As such, one what in which the ASWO could be refined and developed further to allow for deeper semantic annotation, would be to further subclass the current ASWO nodes (such as Rhetorical Support) to account for event more specific forms of these tactics (such as Sympathy, Humour, etc.), and allow more like-for-like comparisons.

There was, overall, a trend of very low reported engagement (Rpl, S, V, Rpr) from participants, across almost all responses. While there was some variation between responses to replying and voting, when asked how likely they would be to share (Table E.7) or report (Table E.9) a post, participants were in fairly close agreement ($\sigma < 1$), stating that they would be less likely to engage than average. The reasons behind the low overall engagement are examined further in Section 6.3.6.6.

Table 6.7: Mean rating for each question, compared with annotations present

	Mean Response to Question								
Annotation	*Co*	*Cr*	*P*	*E*	*O*	*Rpl*	*S*	*V*	*Rpr*
Information	3.475	2.756	3.196	2.678	2.556	2.322	2.045	2.497	1.986
Transition	3.750	2.600	2.825	2.650	2.575	2.525	2.025	2.450	1.900
Logical Attack	3.308	2.966	3.329	2.158	2.301	2.390	2.048	2.548	1.973
Logical Support	3.778	2.815	3.000	2.704	2.963	2.185	2.000	2.333	2.074
Rhetorical Attack	3.527	2.418	2.965	3.094	3.164	2.199	1.863	2.426	2.094
Rhetorical Support	3.580	2.547	2.685	3.420	2.309	2.271	2.122	2.530	1.972
Preference	3.000	2.000	3.500	3.500	3.000	2.000	2.000	2.000	2.500
Persona	3.511	2.427	2.886	3.280	2.853	2.238	2.010	2.463	2.042
Group	3.527	2.873	3.273	2.873	2.873	2.200	1.927	2.618	2.182
Audience	3.510	2.356	2.596	3.740	2.817	2.385	2.269	2.625	2.154
Implied Relationship	3.429	2.143	3.143	2.857	3.429	1.571	1.571	2.429	2.714
Implied Belief	3.800	2.333	3.000	2.967	3.333	2.067	1.800	2.667	2.333
Spam/Advertisement	1.000	1.000	1.000	1.000	1.000	1.000	1.000	4.000	5.000
Unknown	2.750	2.000	2.000	2.250	1.500	1.500	1.500	2.250	1.500
None	3.100	1.700	1.600	3.000	1.700	1.700	1.200	1.700	1.600

Table 6.8: Standard deviation from mean for each question, compared with annotations present

	Standard Deviation from Mean Response to Question								
Annotation	*Co*	*Cr*	*P*	*E*	*O*	*Rpl*	*S*	*V*	*Rpr*
Information	1.131	1.069	1.137	1.178	1.152	1.088	0.950	1.162	0.932
Transition	1.112	1.091	1.138	1.216	1.181	1.072	0.961	1.094	0.970
Logical Attack	1.191	1.036	1.211	1.090	1.088	1.161	1.016	1.188	0.958
Logical Support	0.737	1.020	1.089	0.974	1.261	0.862	0.903	0.903	0.979
Rhetorical Attack	1.107	1.043	1.160	1.221	1.141	1.062	0.906	1.190	0.996
Rhetorical Support	1.113	1.084	1.192	1.142	1.104	1.061	1.033	1.149	0.960
Preference	1.000	0.000	0.500	0.500	0.000	0.000	0.000	0.000	0.500
Persona	1.122	1.032	1.177	1.181	1.206	1.052	0.970	1.167	0.999
Group	1.126	1.113	1.242	1.207	1.192	1.051	0.912	1.168	1.011
Audience	1.109	1.028	1.043	0.971	1.116	1.059	1.058	1.145	0.948
Implied Relationship	0.728	0.639	1.245	0.833	1.178	0.728	0.728	1.498	1.385
Implied Belief	1.108	1.135	1.155	1.303	1.135	0.892	0.792	1.164	1.043
Spam/Advertisement	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Unknown	1.479	1.000	1.000	0.829	0.866	0.866	0.866	1.299	0.866
None	1.578	1.100	0.917	1.483	0.900	1.100	0.600	1.100	1.020

Table 6.9: Average agreement with the statement *This comment is coherent/easy to understand*, grouped by Logic and Rhetoric

Annotation	**Mean**	σ	**t-value**	**p-value**	**U-value**	**p-value**
Logic	3.374	1.150	-1.90	(0.05790)	30830.00	(0.04485)
Rhetoric	3.572	1.107				

Table 6.10: Average agreement with the statement *This comment contains (or appears to contain) credible information*, grouped by Logic and Rhetoric

Annotation	**Mean**	σ	**t-value**	**p-value**	**U-value**	**p-value**
Logic	2.959	1.028	5.30	(< 0.00001)	42980.50	(< 0.00001)
Rhetoric	2.455	1.060				

Table 6.11: Average agreement with the statement *This comment makes (or attempts to make) a persuasive argument*, grouped by Logic and Rhetoric

Annotation	**Mean**	σ	**t-value**	**p-value**	**U-value**	**p-value**
Logic	3.298	1.189	4.14	(0.00005)	41540.00	(0.00003)
Rhetoric	2.848	1.191				

6.3.4 Logic/Rhetoric

This section, and Tables 6.9-6.17, examine how perception of arguments differ purely with regards to logical and rhetorical classifications (both attack and support), and the significance of differences between them. All other classifications are ignored for these tables.

To determine the significance of the difference between participants' responses to comments containing logic and comments containing rhetoric, Welch's t-test (to account for the unequal sample sizes and variance). Given the size of the samples, Welch's t-test should be robust to non-normality. However, in addition to this, a Mann-Whitney U-test (which does not rely on assumptions of normality) was also carried out.

The results show that the features of each comment that differ with significance ($p < 0.01$) are credibility, persuasiveness, entertainment, and offensiveness. The largest difference was in response to whether the post was considered entertaining, with people considering rhetoric to be more entertaining ($|\bar{x_1} - \bar{x_2}| = 0.967$) which was below, though close to, the standard deviation for each. However, the majority of responses across both logic and rhetoric were very similar ($\bar{x_1} - \bar{x_2}| < 0.5$). As mentioned in Section 6.3.3, these homogeneous results may be partly due to the conflict between the inherent subjectivity of different posts. This may also be compounded by the use of logic and rhetoric as such broad categories; therefore, in the following section, these areas are further divided by support and attack, to observe any emerging nuances in perception and engagement.

Table 6.12: Average agreement with the statement *This comment is (or attempts to be) entertaining*, grouped by Logic and Rhetoric

Annotation	Mean	σ	t-value	p-value	U-value	p-value
Logic	2.228	1.087	-9.38	(< 0.00001)	19368.50	(< 0.00001)
Rhetoric	3.195	1.211				

Table 6.13: Average agreement with the statement *This comment is (or attempts to be) offensive*, grouped by Logic and Rhetoric

Annotation	Mean	σ	t-value	p-value	U-value	p-value
Logic	2.404	1.137	-4.01	(0.00007)	27357.00	(0.00010)
Rhetoric	2.830	1.215				

Table 6.14: Average response to the question *Would you be more or less likely to reply to this comment than average?*, grouped by Logic and Rhetoric

Annotation	Mean	σ	t-value	p-value	U-value	p-value
Logic	2.368	1.123	1.67	(0.09651)	36994.00	(0.10829)
Rhetoric	2.200	1.054				

Table 6.15: Average response to the question *Would you be more or less likely to share this comment (to friends/followers/etc.) than average?*, grouped by Logic and Rhetoric

Annotation	Mean	σ	t-value	p-value	U-value	p-value
Logic	2.047	1.002	1.21	(0.22789)	36147.00	(0.25452)
Rhetoric	1.938	0.953				

Table 6.16: Average response to the question *Would you be more or less likely to up-/down-vote this comment than average?*, grouped by Logic and Rhetoric

Annotation	Mean	σ	t-value	p-value	U-value	p-value
Logic	2.520	1.151	0.45	(0.65236)	35092.00	(0.61070)
Rhetoric	2.473	1.185				

Table 6.17: Average response to the question *Would you be more or less likely to report this comment than average?*, grouped by Logic and Rhetoric

Annotation	Mean	σ	t-value	p-value	U-value	p-value
Logic	1.994	0.964	-0.57	(0.56785)	33297.00	(0.59802)
Rhetoric	2.045	0.986				

Table 6.18: Average agreement with the statement *This comment is coherent/easy to understand*, grouped by support and attack

Annotation	Mean	σ	t-value	(p-value)	U-value	(p-value)
Support	**3.603**	**1.082**				
Logical Support	3.778	0.737	1.19	(0.24176)	2545.50	(0.70156)
Rhetorical Support	3.580	1.113				
Attack	**3.445**	**1.140**				
Logical Attack	3.308	1.191	−1.81	(0.07073)	16816.50	(0.07471)
Rhetorical Attack	3.527	1.107				

6.3.5 Support/Attack

As noted in Section 6.3.4, there are different purposes for using logic or rhetoric that may account of the homogeneous results observed. By further breaking the answers down into different types of logic and rhetoric (specifically, support and attack), Tables 6.18-6.26 describe how perception varies further within these classifications. As before, Welch's t-test and a Mann-Whitney U-test were carried out to determine significance, this time comparing logical support directly with rhetorical support, and logical attack directly with rhetorical attack.

Once more, despite separating tactics by support and attack, the majority of responses are relatively close ($|\bar{x_1} - \bar{x_2}| < 0.5$), and again the exception to this was whether posts were considered entertaining and, to a lesser degree, offensive. Supportive posts were considered to be more entertaining ($|\bar{x_1} - \bar{x_2}| = 0.716$) when using rhetorical devices rather than logical devices. Attacks showed the most features with a significant difference in perception, and with an even greater difference in opinion, with rhetorical posts being considered both more entertaining ($|\bar{x_1} - \bar{x_2}| = 0.936$) and more offensive ($|\bar{x_1} - \bar{x_2}| = 0.863$) than logical posts.

What is again clear to see from the result of the significance tests is that while the reported potential engagement (replying, sharing, etc.) when viewing posts is consistently low, it has not been shown to significantly differ when comparing logical or rhetorical tactics, even taking into consideration the differences between supportive and attacking posts. This suggests that the audience of a discussion does not seem to distinguish between the uses of these tactics (at least from the perspective of their own likely engagement with the topic), which goes towards solidifying the suggestion that both modes of argumentation (logical and rhetorical) share equal importance in the minds of viewers.

Table 6.19: Average agreement with the statement *This comment contains (or appears to contain) credible information*, grouped by support and attack

Annotation	Mean	σ	t-value	(p-value)	U-value	(p-value)
Support	**2.564**	**1.081**				
Logical Support	2.815	1.020	1.24	(0.22255)	2739.00	(0.29432)
Rhetorical Support	2.547	1.084				
Attack	**2.613**	**1.074**				
Logical Attack	2.966	1.036	5.07	(< 0.00001)	23956.50	(< 0.00001)
Rhetorical Attack	2.418	1.043				

Table 6.20: Average agreement with the statement *This comment makes (or attempts to make) a persuasive argument*, grouped by support and attack

Annotation	Mean	σ	t-value	(p-value)	U-value	(p-value)
Support	**2.701**	**1.181**				
Logical Support	3.000	1.089	1.36	(0.18180)	2848.00	(0.15407)
Rhetorical Support	2.685	1.192				
Attack	**3.082**	**1.201**				
Logical Attack	3.329	1.211	2.93	(0.00363)	22029.50	(0.00193)
Rhetorical Attack	2.965	1.160				

Table 6.21: Average agreement with the statement *This comment is (or attempts to be) entertaining*, grouped by support and attack

Annotation	Mean	σ	t-value	(p-value)	U-value	(p-value)
Support	**3.328**	**1.148**				
Logical Support	2.704	0.974	-3.42	(0.00152)	1504.50	(0.00079)
Rhetorical Support	3.420	1.142				
Attack	**2.753**	**1.257**				
Logical Attack	2.158	1.090	-7.90	(< 0.00001)	10862.00	(< 0.00001)
Rhetorical Attack	3.094	1.221				

Table 6.22: Average agreement with the statement *This comment is (or attempts to be) offensive*, grouped by support and attack

Annotation	Mean	σ	t-value	(p-value)	U-value	(p-value)
Support	**2.363**	**1.127**				
Logical Support	2.963	1.261	2.51	(0.01745)	3167.50	(0.01038)
Rhetorical Support	2.309	1.104				
Attack	**2.839**	**1.189**				
Logical Attack	2.301	1.088	-7.49	(< 0.00001)	11091.00	(< 0.00001)
Rhetorical Attack	3.164	1.141				

Table 6.23: Average response to the question *Would you be more or less likely to reply to this comment than average?*, grouped by support and attack

Annotation	**Mean**	σ	**t-value**	**(p-value)**	**U-value**	**(p-value)**
Support	**2.265**	**1.038**				
Logical Support	2.185	0.862	−0.46	(0.64941)	2374.50	(0.80720)
Rhetorical Support	2.271	1.061				
Attack	**2.295**	**1.106**				
Logical Attack	2.390	1.161	1.63	(0.10372)	20345.00	(0.12516)
Rhetorical Attack	2.199	1.062				

Table 6.24: Average response to the question *Would you be more or less likely to share this comment (to friends/followers/etc.) than average?*, grouped by support and attack

Annotation	**Mean**	σ	**t-value**	**(p-value)**	**U-value**	**(p-value)**
Support	**2.108**	**1.019**				
Logical Support	2.000	0.903	−0.63	(0.53288)	2311.50	(0.63629)
Rhetorical Support	2.122	1.033				
Attack	**1.953**	**0.956**				
Logical Attack	2.048	1.016	1.82	(0.07042)	20421.00	(0.10106)
Rhetorical Attack	1.863	0.906				

Table 6.25: Average response to the question *Would you be more or less likely to up-/down-vote this comment than average?*, grouped by support and attack

Annotation	**Mean**	σ	**t-value**	**(p-value)**	**U-value**	**(p-value)**
Support	**2.505**	**1.122**				
Logical Support	2.333	0.903	−1.00	(0.32249)	2221.00	(0.43181)
Rhetorical Support	2.530	1.149				
Attack	**2.479**	**1.175**				
Logical Attack	2.548	1.188	0.99	(0.32401)	19820.50	(0.29744)
Rhetorical Attack	2.426	1.190				

Table 6.26: Average response to the question *Would you be more or less likely to report this comment than average?*, grouped by support and attack

Annotation	**Mean**	σ	**t-value**	**(p-value)**	**U-value**	**(p-value)**
Support	**1.975**	**0.957**				
Logical Support	2.074	0.979	0.50	(0.62274)	2588.00	(0.60141)
Rhetorical Support	1.972	0.960				
Attack	**2.050**	**0.968**				
Logical Attack	1.973	0.958	−1.20	(0.23150)	17446.50	(0.24292)
Rhetorical Attack	2.094	0.996				

6.3.6 Rationale

In addition to the Likert-scale questions, participants were provided a free-text area, allowing them to optionally describe the reasons behind the choices they made. Below, a selection of these responses is examined, alongside relevant posts, and the scores given.

6.3.6.1 Coherence

As might be expected, the posts with low scores for coherence ($Co \leq 2$) generally suffered from poor spelling or grammar:

> **Post_10153404551772217_10153405071857217**
> **let him also rule change in presidenny can also change som thing as in develpment.**
> (Information)
>
> ---
>
> (P52) *"I do not understand what the post is trying to say"*

> **Post_10153404551772217_10153404698592217**
> **if him win Clinton don't have to run against him. that no mek sense. when the election every body just go vote.**
> (Information)
>
> ---
>
> (P18, $Co = 1$) *"this comment isn't making any sense"*
>
> ---
>
> (P19, $Co = 2$) *"Poorly written. Makes it hard to understand the point..."*

However, in addition to this, several participants also highlighted a lack of familiarity with the topic area, or a lack of context (with P87 specifically referencing that the acronyms hindered their understanding), as a reason for the post lacking in coherence:

> **Post_10153404015402217_10153404056262217**
> **Money goes to money**
> (Information)
>
> ---
>
> (P54, $Co = 2$) *"Depends on context, not enough information to make it understandable."*

Post_d07pawu
The FSA is also irrelevant. Nusra and/or ISIS would be be the rulers if Assad collapsed. And also, NO the "FSA" is mostly islamists, most of the "secular rebels" have switched around and have become part of the SDF which is part of the YPG.
(Information, Logical Attack)

(P18, $Co = 2$) *"Don't really know the issue discussed in this thread..."*

(P54, $Co = 2$) *"I don't know enough about the topic in the post..."*

(P83, $Co = 2$) *"I think I need more context."*

(P87, $Co = 1$) *"I have no idea what all of the acronyms are, hence have no idea what the target post is actually saying. "*

In terms of posts rated as having high coherence ($Co \geq 4$), the few participants made little direct reference in their reasoning, and those that did were fairly self-explanatory:

Post_700519469144080385
@cnni @JebBush You mean... create tempting targets inside of Syria.
(Information)

(P83, $Co = 5$) *"Comment seems to make sense."*

However, some participants highlighted the fact that some posts were comprehensible due to inferred contextual information (and that even some posts that they considered "dumb" can still be reasonably comprehensible):

Post_700522284872695809
@cnni @realDonaldTrump I doubt you care about anyone, they are just objects.
(Entity, Information, Rhetorical Attack)

(P82, $Co = 4$) *"Seems accurate! (Obviously, if interpreting by inserting an elliptically-elided "to you" at the end.)"*

Post_700456019894345728
@BBCWorld: VIDEO: Star Trek's Shatner opens up about Nimoy https://t.co/rmRlihi5M5

missing Spock #LLAP
(Information)

(P82, $Co = 4$) *"This makes sense, assuming it was posted after Mr. Nimoy's passing; if not, I don't get it at all."*

Post_700301587776983040
@cnni Obama to invade Russia
(Information)

(P82, $Co = 4$) *"One of the dumbest possible comments I can imagine, without descending into total gibberish."*

6.3.6.2 Credibility

The rationale given for posts that received low scores for credibility ($Cr \leq 2$) was generally to do with a lack of provided evidence:

Post_10153404551772217_10153404561637217
Trump is winning evangelicals, moderates, northeasterners, southerners, westerners and his rival democrats are still attacking each other. Barring something completely unforeseeable, Trump will be the next nominee lol
(Information)

(P75, $Cr = 2$) *"Unevidenced opinion."*

Post_d090z15
Let them go. I'm tired of the Scots whining.

There is no utopia South of the border that they are being denied.

Put up a border enforce passport control and let them be governed from Berlin by Merkel.

I'm sure Merkel will look after them.

It's a pity the rest of the UK doesn't get a vote, they would definitely be out.

(Information, Rhetorical Attack)

(P54, $Cr = 1$) *"Well enunciated, but a totally subjective opinion."*

Conversely, posts that provided some verifiable information were considered more credible ($Cr \geq 4$):

Post_10153404551772217_10153404839977217
"It is the absolute right of the State to supervise the formation of public opinion."
Joseph Goebbels
(Entity, Information, Rhetorical Attack)

(P18, $Cr = 4$) *"Seems like a credible argument by using a quote, but doesn't seem to be relevant with the original post"*

Post_d0895b0
Yeah, india supported them during western sanctions.
(Information, Logical Support)

(P19, $Cr = 5$) *"...it doesn't make a particularly strong point (though it does appear to provide additional information)"*

Other credible posts included those that put forth an opinion (but were direct about framing it as such), or those that could be argued for (or against):

Post_d07stsh
Right. And building those pipelines is why the USA and Saudi Arabia started the Syrian war. So it's not like this is just Russia being a dick. Everyone involved in this proxy war has blood on their hands.
(Information, Logical Attack)

(P19, $Cr = 4$) *"The comment has facts that can be argued for or against - so I'd be more likely to interact with this comment."*

Post_d093mo2
I got in a world of trouble for saying something similar on reddit yesterday. I do share this apparently unpopular point of view.

I hope the region can find peace some day but I doubt it will happen with him.
(Information)

(P19, $Cr = 4$) *"I think it's easy enough to understand and they're talking about their own opinions so it is credible..."*

The credibility of certain posts also appeared to be influenced by participants' contextual knowledge (or personal views):

Post_d092ozx
Politics pretty much influences Eurovision's points system though, even if the songs themselves are not political.
(Information, Logical Attack)

(P14, $Cr = 4$) *"Eurovision is influenced by politics is pretty credible"*

6.3.6.3 Persuasion

Posts considered unpersuasive ($P \leq 2$) were generally those adding little real content to the discussion:

Post_700781576624480257
@MailOnline wow thats so fucking original
(Information, Rhetorical Attack)

(P19, $P = 2$) *"...It doesn't contain any real information or attempt to persuade..."*

Post_1184401004952979_1184974498228963
Y'all are funny. Getting mad over nothin
(Group, Rhetorical Attack)

(P19, $P = 1$) *"It's not making any point and without context I don't know who it's even targeted at"*

It is interesting to note that participants often remarked that many of the posts they had considered as *attempting* to persuade ($P \geq 4$), they did not find to be personally persuasive:

Post_10153404551772217_10153404610142217
He d candidate 2 beat at the final if anyone can. Americans need such rugged guy now to help build that great nation again whkich all of us all over d world depended greatly on for provision and protection. Go on Trump you're d best.
(Entity, Information, Rhetorical Support)

(P87, $P = 4$) *"Key word: comment "attempts" to make a persuasive argument. Doesn't succeed though."*

Post_10153404551772217_10153404839977217
"It is the absolute right of the State to supervise the formation of public opinion."
Joseph Goebbels
(Entity, Information, Rhetorical Attack)

(P19, $P = 5$) *'...they're clearly trying to make a point by comparing someone to Nazis. I wouldn't be likely to interact with this."*

Those that the participant did find personally persuasive were often something they already agreed with prior to seeing the post, or something that they felt could be engaged with:

Post_700309701485932544
@cnni This country don't grow up & stop playing cowboy were going to be the ones torn up from within! #FeelTheBern #WorkingClass #NoMoreWar
(Entity, Information, Rhetorical Support)

(P18, $P = 4$) *"Agree with target post"*

Post_d08ganb
How quickly can a referendum be called anyways?
(Transition)

(P18, $P = 4$) *"Good argument: if it's takes a long time to be able to call a referendum, then the threat/warning is kind of pointless"*

Post_700711259688165376
@SkyNews @Beyoncé She sort of loses any respect dancing around in her near on underwear. Stay out of politics, or do it correctly.
(Entity, Information, Rhetorical Attack)

(P18, $P = 4$) *"Not sure if I 100% agree with the target post (don't think it's wrong for her to be involved in politics), or what beyonce did in superbowl, but it does make some good arguments: protesting political issues should be done correctly, "*

6.3.6.4 Entertainment

The primary reason given for judging a post as attempting to be entertaining ($E \geq 4$) concerned the use of humour, including crass or vulgar humour (even if they did not personally find it funny):

Post_d087izl
"There sure are alot of people mad about buckets today, sir."
(Audience, Entity, Rhetorical Support)

(P54, $E = 4$) *"The mental picture that makes the comment humorous wasn't immediately apparent, though it was immediately obvious it was attempted humour."*

Post_10153404551772217_10153404572497217
Never trust a political aspirant whose tie tip reaches his dick.
(Entity, Rhetorical Attack)

(P45, $E = 5$) *"It is emotionally engaging so prompts replying. It is entertaining that promots sharing. "*

Post_d087me0
> **This guy saved Syria the same way dried scabs from lepers enhances the flavor of ice cream**.

You watch your mouth god damn it. I happen to know plenty of good natured dried scabs and your demeaning attitude is hurtful.
(Audience, Entity, Humour, Rhetorical Support)

(P54, $E = 4$) *"Attempted humor, not funny."*

Post_10153404551772217_10153404579852217
Soon jolly ol England will have to reclaim these colonies [guardsman-emoji]
(Audience, Entity, Information, Rhetorical Attack, Rhetorical Support)

(P82, $O = 4$) *"It's mildly offensive, but mostly it's just a bad, totally pointless joke, best ignored."*

(P83, $O = 4$) *"Just a fun comment."*

6.3.6.5 Offence

While participants rarely highlighted reasons for why they found a post inoffensive, they were much more clear about what about posts they considered offensive ($O \geq 4$). Participants felt, perhaps unsurprisingly, that posts that contained swearing or foul language were generally offensive (although this was somewhat subjective), or those that directly insulted a person (whether within the discussion, or the topic of it), and typically stated that they would not engage with it:

Post_700781576624480257
@MailOnline wow thats so fucking original
(Information, Rhetorical Attack)

(P19, $O = 4$) *"It's just swearing so not particularly offensive..."*

Post_700697066943569920
@skynews @mikkil She is black. So why don't they just shoot her instead???
(Group, Information, Rhetorical Attack)

(P75, $O = 4$) *"I don't engage with racial hatred discussions. There's no rational discussion."*

Post_700375759798607872
@BBCNews

of course the video about skinny models is fronted by a jealous lesbian dressed as a man

Lefty bullshit
(Entity, Group, Information, Rhetorical Attack)

(P86, $O = 5$) *"I typically ignore this type of stuff."*

6.3.6.6 Engagement

Due to the overall low-levels of reported potential engagement, the remaining categories (Rpl, S, V, and Rpr) are grouped together.

Participants had different opinions on how emotional language would change their behaviour; one explained that they were more likely to reply to posts that didn't seem emotionally charged, whereas another felt the opposite.

Post_d07pawu
The FSA is also irrelevant. Nusra and/or ISIS would be be the rulers if Assad collapsed. And also, NO the "FSA" is mostly islamists, most of the "secular rebels" have switched around and have become part of the SDF which is part of the YPG.
(Information, Logical Attack)

(P45, $Rpl = 4, S = 3, V = 3, Rpr = 2$) *"I liked the non-emotional tone. So I would comfortable replying. But because my emotions are not engaged, I am less inclined to share..."*

Post_700685814045474818
@MailOnline death [skull-emoji] for the unnecessary and overuse of the character in heavy metal band names
(Audience, Entity, Humour, Rhetorical Support)

(P82, $Rpl = 4, S = 2, V = 4, Rpr = 2$) *"People are facing execution, and someone posts a dumb joke? It would be pretty funny in other contexts, but this is gross. I'd be more likely to reply just to call them out for being an ass."*

Post_d08ganb
How quickly can a referendum be called anyways?
(Transition)

(P19, $Rpl = 4, S = 1, V = 1, Rpr = 1$) *"It asks a question. I'd be more likely to reply to answer it."*

Post_10153404551772217_10153404572497217
Never trust a political aspirant whose tie tip reaches his dick.
(Entity, Rhetorical Attack)

(P45, $Rpl = 4, S = 4, V = 2, Rpr = 4$) *"It is emotionally engaging so prompts replying. It is entertaining that promots* [sic] *sharing."*

There was also a relatively consistent consensus that participants were more likely to find a post persuasive, or engage with it, if they already personally agreed with the content (although some participants reported a deliberate engagement with content they disagreed with):

Post_d08kmf5
***FPTP, not the most representative of democratic systems. Not really that democratic at all once you start thinking about it.**
(Information, Logical Attack)

(P87, $Rpl = 4, S = 3, V = 5, Rpr = 1$) *"The bias is: I am much more likely to share/upvote a comment if I agree with its contents."*

Post_1184401004952979_1184974498228963
Y'all are funny. Getting mad over nothin
(Group, Rhetorical Attack)

(P18, $Rpl = 1, S = 2, V = 5, Rpr = 1$) *"...Since I agree with the comment, I'm likely for me to vote up, but unlikely to comment further because there's nothing more to discuss on this issue."*

Post_d07pawu
The FSA is also irrelevant. Nusra and/or ISIS would be be the rulers if Assad collapsed. And also, NO the "FSA" is mostly islamists, most of the "secular rebels" have switched around and have become part of the SDF which is part of the YPG.
(Information, Logical Attack)

(P19, $Rpl = 3, S = 3, V = 3, Rpr = 1$) *"...I tend to upvote content that I agree with and share content I disagree with..."*

As might be expected, participants were more likely to report posts that did not appear to be entering the discussion in good faith, whether through insults or derailing the topic.

Post_d08vh7c
https://www.reddit.com/r/isrconspiracyracist/search?q=Dhylan&restrict_sr=on
(Information)

(P75, $Rpl = 1, S = 1, V = 5, Rpr = 4$) *"Baiting."*

Post_d095om7
true

true

I lol'd but true nevertheless.
(Logical Support, Rhetorical Support)

(P75, $Rpl = 1, S = 1, V = 4, Rpr = 4$) *"Point #3 indicates the respondent isn't likely to engage in polite debate."*

Post_700713867060576258
@SkyNews @mynameisbrogs_
(None)

(P75, $Rpl = 1, S = 1, V = 4, Rpr = 4$) *"Appears to be spam"*

Several of the rationales given justified the low engagement scores given ($Rpl \leq 3, S \leq 3, V \leq 3, Rpr \leq 3$). These were broadly in two camps: either due a general disinterest in the subject at hand, or due to the post being unclear or not credible.

Post_d08zyyr
What happened to the days of installing puppet governments?

Now they just leave voids, to be filled with terror groups.

Ooooooohhhhh.
(Information)

(P51, $Rpl = 3, S = 3, V = 3, Rpr = 3$) *"I don't really reply to comments on social media, but do often read them. Hence my 'neutral' more/less likely answers to these questions."*

Post_d07y4s0
Saved? Mate, you got them into this mess in the first place and wallowed

around for years before the Russians showed up to straighten things out.
(Information, Logical Attack)

(P86, $Rpl = 1, S = 1, V = 1, Rpr = 1$) *"I just don't care about politics."*

Post_10153404015402217_10153404031222217
Why don't poor people win the lottery? Middle class lottery perhaps
(Information, Transition)

(P19, $Rpl = 1, S = 1, V = 1, Rpr = 1$) *"...I'm not particularly interested so I wouldn't be likely to engage them."*

Post_700758214728613888
@SkyNews @thecampaignbook He was in that lift all day talking to people, i bet it fucking wreaked of shit, they couldnt wait 2 get out i bet
(Entity, Information, Rhetorical Attack)

(P48, $Rpl = 1, S = 1, V = 1, Rpr = 1$) *"Not clear the respondent's intention"*

Post_d0895b0
Yeah, india supported them during western sanctions.
(Information, Logical Support)

(P18, $Rpl = 2, S = 2, V = 2, Rpr = 1$) *"Don't know how credible this information is, so I wouldn't interact with this comment"*

Conversely, some participants explained their enthusiasm for interacting with certain comments particularly *because* of this.

Post_10153404015402217_10153404031222217
Why don't poor people win the lottery? Middle class lottery perhaps
(Information, Transition)

(P18, $Rpl = 4, S = 4, V = 4, Rpr = 1$) *"This is a stupid argument, so I'm likely to interact with it."*

Others pointed out they were more likely to interact with posts that appeared to have a central conclusion that could actually be argued for or against.

Post_d07stsh

Right. And building those pipelines is why the USA and Saudi Arabia started the Syrian war. So it's not like this is just Russia being a dick. Everyone involved in this proxy war has blood on their hands.

(Information, Logical Attack)

(P19, $Rpl = 4, S = 1, V = 4, Rpr = 1$) *"The comment has facts that can be argued for or against - so I'd be more likely to interact with this comment."*

Post_d092dtj

> **It's a historical event that fundamentally changed the country.**

Which event? Deportation of Tatars under Joseph Stalin didn't fundamentally change USSR.

(Information, Logical Attack, Transition)

(P19, $Rpl = 4, S = 2, V = 3, Rpr = 1$) *"I don't know a lot about the event but I would be more likely to respond to this as there is a clear point that could be discussed..."*

6.4 Summary

An experiment was carried out in which 60 participants were shown 436 unique comments from 3 different platforms of the social web. Each participant was asked to rate 20 comments on how they perceived it (in terms of coherence, credibility, persuasiveness, entertainment and offence) and how they would engage with it (in terms of replying, sharing, voting on or reporting to moderators). In total, 746 comments were presented (giving 6714 answered questions).

Due to a relatively high standard deviation of responses between participants, the results point to a degree of variation between participants' answers. This is likely due to a combination of factors, in particular the individual variance of the posts (in terms of tone, implicit meaning, context, etc.), and the natural subjectivity with which people view argument and how they respond.

The clearest observable difference in significance between Logic and Rhetoric arose when considering support and attack separately. Rhetorical Support was considered more entertaining than Logical Support ($|\bar{x_1} - \bar{x_2}| = 0.716$, $U = 1504.50$, $p = 0.00079$), and Rhetorical Attacks were found to be more entertaining ($|\bar{x_1} - \bar{x_2}| = 0.936$, $U = 10862.00$, $p < 0.00001$) and more offensive ($|\bar{x_1} - \bar{x_2}| = 0.863$, $U = 11091.00$, $p < 0.00001$) than Logical Attacks. Other observed significant differences had a lower degree of difference between means, and different categories of engagement showed no significant differences at all.

In answer to research question *R*4 (*Do rhetorical techniques affect the way in which users perceive and engage with online argument?*), these results partly support the experimental hypothesis proposed in Section 6.1 (and the greater hypothesis presented in Section 1.2), suggesting that in some respects, users do see rhetorical techniques as distinct from logical techniques (often considering them to be more entertaining and/or offensive). However, in terms of the the general perception of argumentation, they do not appear to be as significantly different as initially thought, as well as appearing very similar in terms of users' reported potential engagement. While an unexpected development, this does perhaps lend further support to the proposal that rhetorical argumentation has a natural place in discussions on the social web, as users do not discard it out-of-hand when considering an on-going debate. Therefore, it is still fair to consider rhetorical contributions to be a valuable form of argumentation and—as a result—to model it alongside logical forms.

Chapter 7

Conclusions and Future Work

7.1 Summary of Work and Findings

The primary goal of this thesis has been to examine eristic argumentation on the social web (with particular attention to rhetorical tactics applied within, and how these are perceived by those that witness them), presenting the hypothesis that rhetorical tactics—humorous, abusive or otherwise—are an important factor of social argumentation, and should be included in formal models of argument.

Chapter 3 presented a preliminary investigation into the capabilities of current models of social argumentation, in which a small case-study was carried out, examining three different areas of the social web to determine the strengths and weaknesses of modelling social, eristic argument on the web using current techniques. This showed that, firstly, and most importantly, rhetorical tactics are shown to be present throughout the argumentation in each of the case studies. Secondly, this preliminary work indicated that existing techniques for modelling argumentation were insufficient to capture the structure and dynamic of argumentation taking place on the social web.

Chapter 4 proposed an extension to the existing AIF ontology to specifically take into account the application of rhetorical support and attack as used on the social web. The previous case-study was repeated in greater detail, using these new proposals to catalogue the additional examples of rhetorical tactics used, examine their prevalence in web-based argumentation and their correlation with machine readable features (such as post length, language, etc.). These proposals were then augmented further to take into account the role of audience, and the ability of participants to imply information and/or relationships without explicitly stating them. These were validated by means of an expert review; experts from a variety of domains including argumentation modelling, the social web, linked data, and philosophy were consulted on the benefits, drawbacks, and potential future direction of the Argumentation on the Social Web Ontology. These experts reacted favourably to many of the additions made and, based on their feedback, further clarification of the ASWO was made.

Chapter 5 detailed additional data collection and annotation on a much larger scale that previously. This data was analysed structurally in terms of social structures (including average comment length, average number of replies, etc.) and in terms of the annotated classifications. Three individual threads were then examined in detail, showcasing the different types of argumentation that it is possible to observe across different social media platforms. This data was then used in Chapter 6, forming the basis of an experiment into how the application of logical and rhetorical comments affect the perception, and reaction to, a discussion on social media. 60 participants were shown 436 unique comments from the data sample, with 746 comments in total presented (giving 6714 answered questions). These results were analysed individually for each of the nine questions participants were asked per comment with regard to each individual classification, logic and rhetoric, and support and attack.

The primary limitation of this work is the necessity to manually annotate all the data using the ASWO (even factoring the affordances of using a simpler classification system over a full relationship model). This is time consuming and subjective, but as yet there is no way to circumvent this process and automatically extract premises and conclusions, particularly given the informal language patterns commonly found on social media. A further constraint is that only English-language sites are examined. There are, of course, many other social media services that cater to audiences of different languages, such as *Renren*[1] for China or *VKontakte*[2] for Eastern Europe. However, this separation is mitigated by the fact that different languages (and different cultures) have their own rhetorical structures and argumentation schemes (Van Eemeren and Grootendorst, 2004, p. 21). As a result, attempting to analyse multiple sites with different primary languages concurrently would distort any patterns that might emerge in the argument structure of the users.

7.1.1 Hypothesis and Research Questions

Revisiting the hypothesis initially proposed in Section 1.2:

"A model of eristic argumentation on the social web should include both logical and rhetorical tactics, as the inclusion of rhetorical techniques affects the way in which users perceive and engage with the argument"

This was resolved into four distinct research questions:

R1. *Is modelling eristic argumentation a valuable direction of work?*

R2. *Are current frameworks and tools sufficient to model eristic argumentation on the social web?*

R3. *How should rhetorical techniques be included in a model of eristic argumentation on the social web?*

[1] http://renren.com/
[2] http://vk.com/

R4. *Do rhetorical techniques affect the way in which users perceive and engage with the argument?*

Based on the proceeding body of work, these questions can now effectively be answered as follows:

*R*1: The question as to whether modelling eristic argumentation can be considered valuable is answered satisfactorily enough through the observation of current literature in Chapter 2, with existing work suggesting that eristic argument is indeed present on the social web (Schneider et al., 2012, 2014) and is indeed worth modelling, particularly with regards to the notions of audience (Berland and Forte, 2010) and anti-social behaviour (Papacharissi, 2004; Sood et al., 2012; Jane, 2014). This view is strengthened by examination of existing behaviours on the social web in Chapters 3 and 4, by observing the prevalence and impact of eristic behaviours within ongoing argument, and is then further confirmed through expert review of the Argumentation on the Social Web Ontology, with experts stating that: *"...if we're going to have a realistic model of how people argue, we've got to look at how people really argue rather than how our "ideal reasoner" would argue"* and *"I think modelling social argumentation is very important...I want to say it's useful in trying to help people argue 'better'."*

*R*2: The question as to whether current models of argument are sufficient in capturing eristic argumentation is answered through the preliminary work carried out in Chapter 3. In short: they are not (albeit by design). The majority of formalised argumentation modelling frameworks examined in Chapter 2 make the assumption of a dialectic approach in which all participants are engaging in good faith. This is a deliberate design decision which allows them to produce models that can be reasoned over with the use of intelligent agents. However, in doing so there are many (perhaps "unreasonable") argumentation techniques—particularly those that incorporate social pressures—that are not accounted for when applying such frameworks, such as directed abuse, playing to the crowd, as well as levity and humour. In addition to this, there are the features of discourse that are specific to social media, including social relationships (i.e. the links between profiles), reputation features (such as "Likes" or "Retweets"), and the use of avatars (see Section 7.2). This contribution was presented at the Computational Models of Natural Argument workshop (Blount et al., 2014).

*R*3: The question as to how exactly eristic argumentation, and rhetorical techniques, should be modelled was answered in Chapter 4, through the creation of a series of proposals for an extension to the AIF (partly incorporating the SIOC ontology), dubbed the Argumentation on the Social Web ontology. These proposals were refined through a practical implementation (using the ASWO to model a real example of argumentation on the social web) and expert review, and resulted in a new set of nodes specifically revolving around modelling rhetorical techniques, the notion of a viewing audience, and the social pressures and relationships between them. These additional nodes were as follows: Rhetorical Attack and Rhetorical Support, in which a participant attempts to discredit or uphold a position by using emotion and social pressure, without using evidence or logic (such as *"This sucks"* or *"Amen buddy!"*); Personas, which

form a key part of the model, representing not just a participants social media account, or even the participant as a real-world person, but their inherent character that they assume during the course of the argument, allowing it to be directly supported or attacked in the same manner a logical tactic might support or attack a claim or piece of data; Groups, which define abstract groupings of Personas (such as political affiliation, or pro-/con- sides of the debate); Audience, which models the notion of all Personas actively engaged in the debate, as well as those passively watching; Implied Relationship, in which participants suggest that other Personas may belong to a particular Group (which can be either a positive or a negative sentiment); and finally, Implied Belief, for use when a participant makes the suggestion or allegation that a particular Persona, Group or the Audience itself, believes a certain claim. This framework was used to determine the classification system to build the dataset in Chapter 5, and was presented at ACM Hypertext (Blount et al., 2015a) and the International Conference on Computational Models of Argument (Blount et al., 2016).

*R*4: Finally, the question as to whether the use of logical or rhetorical techniques affects the audiences perception of the argument is answered in Chapter 6, in which an experiment was carried out to show social-media users a series of comments using a variety of argumentation techniques (from the dataset built in Chapter 5) and gauge their responses in terms of both perception and engagement. Perhaps surprisingly, the majority of participants reported a very low engagement across the board, regardless of the type of argument used, strengthening the position that users are more engaged if they purposefully seek out a discussion rather than if they are shown it, or see it by chance. However, rhetorical tactics were perceived as being more entertaining (and in many cases more offensive) than logical tactics, although in all other respects, logical and rhetorical techniques elicited very similar reactions from the participants. While this does only partly supports the initial hypothesis, it nevertheless strengthens the position that—due to the way participants engage with logical and rhetorical tactics on a more or less equal basis—it is important to model *both* these aspects of argumentation.

7.2 Proposals for Future Work

Following the investigations that have been carried out in this thesis, and the findings summarised in Section 7.1, a selection of particular avenues of future work is highlighted. The development of the Argumentation on the Social Web Ontology has been, and should continue to be, an evolving process. Further refinement and expert review will provide an even more robust framework with which to drive further research in the social-argumentation space. Below are several potential paths that could shape the future of ASWO and develop it to suit specific research paths.

7.2.1 Extended Perception Study

One method of extending the work carried out in Chapter 6 is to extend the study by repeating the experiment, but with further context of the overall discussion. That is, showing participants either the full thread or a reasonable-sized subset (for, potentially, a limited amount of time, both for scaling purposes, and to mimic the casual consumption of social media), then asking them to rate their overall impression of it, given the full context, on a set of questions similar to those used in the previous perception experiment (i.e. measuring both thoughts of and reactions to the content). This can then be compared with the results taken from the study of individual comments, to see if further context affects how the audience of a discussion perceives or engages with it.

7.2.2 Social Features and Self-Presentation

While the work presented here has succeeding in capturing certain aspects of social argumentation - namely that it can be a process of catharsis without the need for logic - there are other elements that would likely also valuable to capture. These include features unique to the social web (avatars, reputation systems, etc.). For example, compare a user with the default profile picture and no social links making a post, with a "normal" account, complete with personal profile picture and a reasonable number of links, making the same post. Alternatively, compare the account of prominent celebrity which has many social links making a post with many up-votes to an "official" account of a organisation (relevant to the topic at hand) making the same post with comparatively few up-votes. It is highly likely that these posts will be perceived (and acted on) very differently by the same audience (Zanbaka et al., 2006). Another key element of this is the "Proteus" effect: that a person acting under the guise of an avatar will subtly change their behaviour to meet the mental expectations they have of that particular avatar (Yee and Bailenson, 2007).

One facet of this could be achieved by reproducing an experiment similar to that in Chapter 6, but including (or even spoofing, for the purposes of providing a variety of experimental groups with known constants) additional social features such as avatar, the number of up-votes, replies or shares a particular post has, and so on, to determine any effect these elements have on the audiences perception.

This contribution to the field could then be used to assist further work in a number of other areas, such as another metric for use with adaptive recommendation techniques to match people based on preferred argumentation strategies (Guy et al., 2010), or the development of argumentation frameworks that integrate with the social web (Torroni et al., 2010).

7.2.3 AI and Reasoning

As is the focus of many researchers in this field, attention can be given to the use of artificial intelligence and argumentation, whether by reasoning over a model of argument in an attempt to determine the most valid argument and subsequent course of action (Caminada and Amgoud, 2007) or by using the model to influence the techniques and strategies of intelligent agents involved in dialogue games (Reed et al., 2008). However, the fact that the eristic features of the model are unlikely to be practical (or appropriate) for the use of reasoning, or governing inter-agent negotiations is likely what has caused them to be currently excluded from the majority of formal models. Disregarding this, the weakness of this approach is that the model cannot, at this stage, be automatically constructed, but must be created through a time and labour intensive process of manual annotation. Therefore, using the model as a basis of reasoning over argumentation in general is ultimately flawed. Any gains that were achieved in this area would be rendered moot by the cost of creating a model for every argumentation to be reasoned over, and rendered impractical on a web-scale.

An alternative avenue would be to generate this model from the arguments$_2$ themselves, by means of natural language processing (Palau and Moens, 2009), the use of social machines (Hendler and Berners-Lee, 2010) or some combination thereof. This would go some way towards solving a large outstanding issue in the field (Schneider et al., 2013, p. 31-32). While working towards a means of automatically generating the model has potential, it is likely that the social and eristic nature of the arguments to be modelled is the very thing that hinders this approach. Web-based culture and language is made up of many disparate groups, and continues to rapidly and constantly evolve, which renders current natural language processing impractical in the short term and ineffective in the long term, without the use of domain-specific normalisation techniques that are expensive or inaccurate (Han and Baldwin, 2011; Eisenstein, 2013). While the findings in Section 4.2.2 point towards a means of broadly classifying a post as containing different types of logical or rhetorical elements, with reasonable probability, the overall structure may be difficult to model automatically.

7.2.4 Crowdsourcing

Clearly, at this stage, human input cannot be wholly eliminated from the process of modelling argumentation in this way. However, crowdsourcing has been shown to be a viable method of distributing the annotation required for modelling argumentation at a larger scale (Ghosh et al., 2014). With the use of crowdsourcing or social machines, the large effort cost of annotating arguments$_2$ could be distributed across participants to a manageable level. To further this work, an experiment to determine the prerequisites of a sufficiently competent annotator could be undertaken, examining the effects of both the complexity of the annotation system (fine- or coarse-grained, classification-only or relationship-mapping), and the amount of training non-experts receive, against the performance when compared to expert annotators.

7.3 Final Conclusions

Argumentation, like the social web itself, is a diverse construct that is challenging to model, but has huge potential if correctly harnessed. It has been shown that a failure to accurately model these social argumentation strategies is detrimental to the goal of studying how discussions evolve on the social web. Rhetoric and logic are both important aspects of online social argumentation; to accurately model how arguments occur and evolve across social media, it is important to take into account all the techniques and tactics that are employed.

While it is difficult to determine the objective value of an individual contribution to a discussion (if indeed such a thing can be said to exist), to define all logical contributions (and only logical contributions) as inherently valuable is a naive approach. Being able to accurately record all aspects of argumentation on social media, including those that rely on rhetorical force and social pressures, is the first step towards being able to accurately analyse informal argument on an enormous scale.

The work presented in this report provides a novel framework for modelling eristic and rhetorical argumentation, ideal for use in modelling social argumentation, and demonstrates some of the structures that may be observed when this model is applied to real examples of argument on the social web.

It is hoped that, by bringing rhetorical and logical models of argumentation together with the computational modelling of social media, the Argumentation on the Social Web Ontology has the potential to be a powerful tool in both the understanding of eristic and social argumentation, and the use of social media more generally. By laying the initial groundwork in modelling the common (and fundamental) aspects of eristic argumentation, this raises the possibility of future models that explore these constructs in even greater depth, that could lead the way in supporting the development of new tools that could help communities encourage and moderate argumentation, and counter diverse problems from echo-chambers and groupthink, to trolling, anti-social behaviour, and harassment.

References

Alagoz, E. (2013). Social argumentation in online synchronous communication. *International Journal of Computer-Supported Collaborative Learning*, 8(4):399–426.

Anderson, A., Huttenlocher, D., Kleinberg, J., and Leskovec, J. (2012). Discovering value from community activity on focused question answering sites: a case study of stack overflow. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 850–858. ACM.

Anderson, A. A., Brossard, D., Scheufele, D. A., Xenos, M. A., and Ladwig, P. (2014). The nasty effect: Online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication*, 19(3):373–387.

Arceneaux, K. (2012). Cognitive biases and the strength of political arguments. *American Journal of Political Science*, 56(2):271–285.

Aurisicchio, M. and Bracewell, R. (2013). Capturing an integrated design information space with a diagram-based approach. *Journal of Engineering Design*, 24(6):397–428.

Austin, J. L. (1962). *How to do things with words*. Oxford University Press.

Bailenson, J. N. and Blascovich, J. (2004). Avatars. In *Encyclopedia of Human-Computer Interaction, Berkshire Publishing Group*, pages 64–68. Berkshire Publishing Group.

Bench-Capon, T. J. and Dunne, P. (2002). Value based argumentation frameworks. In *Research Report ULCS-02-001*, pages 444–453. Department of Computer Science, The University of Liverpool.

Bench-Capon, T. J. M. and Dunne, P. E. (2007). Argumentation in artificial intelligence. *Artif. Intell.*, 171(10-15):619–641.

Berland, L. K. and Forte, A. (2010). When students speak, who listens?: constructing audience in classroom argumentation. In *Proceedings of the 9th International Conference of the Learning Sciences-Volume 1*, pages 428–435. International Society of the Learning Sciences.

Bertea, S. (2004). Certainty, reasonableness and argumentation in law. *Argumentation*, 18(4):465–478.

Bex, F., Lawrence, J., Snaith, M., and Reed, C. (2013). Implementing the Argument Web. *Communications of the ACM*, 56(10):66–73.

Bex, F., Snaith, M., Lawrence, J., and Reed, C. (2014). ArguBlogging: An application for the Argument Web. *Web Semantics: Science, Services and Agents on the World Wide Web*, 25:9 – 15.

Blackburn, J. and Kwak, H. (2014). Stfu noob!: predicting crowdsourced decisions on toxic behavior in online games. In *Proceedings of the 23rd international conference on World wide web*, pages 877–888. ACM.

Blount, T., Millard, D., and Weal, M. (2016). An Ontology for Argumentation on the Social Web: Rhetorical Extensions to the AIF. In *International Conference on Computational Models of Argument*.

Blount, T., Millard, D. E., and Weal, M. J. (2014). Towards Modelling Dialectic and Eristic Argumentation on the Social Web. In *14th workshop on Computational Models of Natural Argument*.

Blount, T., Millard, D. E., and Weal, M. J. (2015a). An Investigation into the Use of Logical and Rhetorical Tactics within Eristic Argumentation on the Social Web. In *ACM Conference on Hypertext and Social Media*.

Blount, T., Millard, D. E., and Weal, M. J. (2015b). On the Role of Avatars in Argumentation. In *Proceedings of the 2015 Workshop on Narrative & Hypertext*, pages 17–19. ACM.

Botting, D. (2015). Inferences and illocutions. *Argument & Computation*, 6(3):246–264.

Braet, A. C. (1992). Ethos, pathos and logos in Aristotle's Rhetoric: A re-examination. *Argumentation*, 6(3):309–315.

Breslin, J. G., Decker, S., Harth, A., and Bojars, U. (2006). SIOC: An Approach to Connect Web-Based Communities. *International Journal of Web Based Communities*, 2(2):133–142.

Bucher, T. (2012). A technicity of attention: How software 'makes sense'. *Culture machine*, 13:1–23.

Buckingham Shum, S., Motta, E., and Domingue, J. (2000). Scholonto: an ontology-based digital library server for research documents and discourse. *International Journal on Digital Libraries*, 3(3):237–248.

Budzynska, K., Janier, M., Reed, C., and Saint-Dizier, P. (2016). Theoretical foundations for illocutionary structure parsing 1. *Argument & Computation*, 7(1):91–108.

Budzynska, K. and Reed, C. (2011a). Speech acts of argumentation: Inference anchors and peripheral cues in dialogue. In *Proceedings of the 10th AAAI Conference on Computational Models of Natural Argument*, AAAIWS'11-10, pages 3–10. AAAI Press.

Budzynska, K. and Reed, C. (2011b). Whence inference? *University of Dundee Technical Report*.

Budzynska, K. and Reed, C. (2012). The Structure of Ad Hominem Dialogues. In *Proceedings of Computational Models of Argument*, volume 245, pages 410–421.

Caminada, M. and Amgoud, L. (2007). On the evaluation of argumentation formalisms. *Artificial Intelligence*, 171(5):286 – 310.

Chalamish, M., Gabbay, D., and Schild, U. (2011). Intelligent Evaluation of Evidence Using Wigmore Diagrams. In *Proceedings of the 13th International Conference on Artificial Intelligence and Law*, ICAIL '11, pages 61–65, New York, NY, USA. ACM.

Chalamish, M., Hazoom, M., and Schild, U. (2013). Semi-automatic Creation of Wigmore Diagrams. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law*, ICAIL '13, pages 181–185, New York, NY, USA. ACM.

Chesñevar, C., McGinnis, J., Modgil, S., Rahwan, I., Reed, C., Simari, G., South, M., Vreeswijk, G., and Willmott, S. (2006). Towards an argument interchange format. *Knowledge Engineering Review*, 21(4):293–316.

Conklin, J. and Begeman, M. L. (1987). gibis: A hypertext tool for team design deliberation. In *Proceedings of the ACM Conference on Hypertext*, HYPERTEXT '87, pages 247–251, New York, NY, USA. ACM.

Conti, G. and Sobiesk, E. (2010). Malicious interface design: Exploiting the user. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, pages 271–280, New York, NY, USA. ACM.

Corbett, E. P. and Connors, R. J. (1999). *Classical rhetoric for the modern student*. Oxford University Press.

Davis, J. P. (2002). The experience of 'bad' behavior in online social spaces: A survey of online users. *Social Computing Group, Microsoft Research*.

Dung, P. M. (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial intelligence*, 77(2):321–357.

Dunne, P. E. (2016). I Heard You the First Time: Debate in Cacophonous Surroundings. In *Computational Models of Argument: Proceedings of COMMA 2016*, volume 287, page 287. IOS Press.

Duthie, R., Budzynska, K., and Reed, C. (2016). Mining ethos in political debate. In *COMMA*, pages 299–310.

Eisenstein, J. (2013). What to do about bad language on the internet. In *Human Language Technologies*, pages 359–369. North American Chapter of the Association for Computational Linguistics.

Ghosh, D., Muresan, S., Wacholder, N., Aakhus, M., and Mitsui, M. (2014). Analyzing argumentative discourse units in online interactions. In *Proceedings of the First Workshop on Argumentation Mining*, pages 39–48.

Gilbert, E., Bergstrom, T., and Karahalios, K. (2009). Blogs are echo chambers: Blogs are echo chambers. In *42nd Hawaii International Conference on System Sciences, HICSS'09*, pages 1–10. IEEE.

Goodwin, J. and Fisher, A. (2000). Wigmore's Chart Method. *Informal Logic*, 20(3):223–243.

Guy, I., Zwerdling, N., Ronen, I., Carmel, D., and Uziel, E. (2010). Social media recommendation based on people and tags. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, pages 194–201. ACM.

Hahn, U., Oaksford, M., and Corner, A. (2005). Circular arguments, begging the question and the formalization of argument strength. In *Proceedings of AMKLC'05, International Symposium on Adaptive Models of Knowledge, Language and Cognition*, pages 34–40.

Han, B. and Baldwin, T. (2011). Lexical normalisation of short text messages: Makn sens a #twitter. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, volume 1, pages 368–378. Association for Computational Linguistics.

Hanna, R., Rohm, A., and Crittenden, V. L. (2011). Were all connected: The power of the social media ecosystem. *Business horizons*, 54(3):265–273.

Hendler, J. and Berners-Lee, T. (2010). From the semantic web to social machines: A research challenge for ai on the world wide web. *Artificial Intelligence*, 174(2):156–161.

Hermida, A., Fletcher, F., Korell, D., and Logan, D. (2012). Share, like, recommend. *Journalism Studies*, 13(5-6):815–824.

Herring, S., Job-Sluder, K., Scheckler, R., and Barab, S. (2002). Searching for Safety Online: Managing "Trolling" in a Feminist Forum. *The Information Society*, 18(5):371–384.

Hodson, H. (2013). Moderate your language. *New Scientist*, 218(2912):18.

Hooi, R. and Cho, H. (2013). Deception in avatar-mediated virtual environment. *Computers in Human Behavior*, 29(1):276 – 284.

Jane, E. A. (2014). "your a ugly, whorish, slut": Understanding ebile. *Feminist Media Studies*, 14(4):531–546.

Janier, M., Lawrence, J., and Reed, C. (2014). Ova+: An argument analysis interface. In *Computational Models of Argument: Proceedings of COMMA*, volume 266, page 463.

Jiménez-Aleixandre, M. P. and Erduran, S. (2007). Argumentation in science education: An overview. In *Argumentation in science education*, pages 3–27. Springer.

Jørgensen, C. (1998). Public Debate – An Act of Hostility? *Argumentation*, 12(4):431–443.

Kaplan, A. M. and Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business horizons*, 53(1):59–68.

Karunatillake, N. C., Jennings, N. R., Rahwan, I., and McBurney, P. (2008). Formal Semantics of ABN Framework. Technical report, University of Southampton.

Kennedy, G. A. (1991). *Aristotle on rhetoric: a theory of civic discourses*. Oxford University Press.

Kerferd, G. B. (1981). *The Sophistic Movement*. Cambridge University Press.

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., and Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business horizons*, 54(3):241–251.

Kietzmann, J. H., Silvestre, B. S., McCarthy, I. P., and Pitt, L. F. (2012). Unpacking the social media phenomenon: towards a research agenda. *Journal of Public Affairs*, 12(2):109–119.

Klein, M. (2010). Using metrics to enable large-scale deliberation. In *Proceedings of Collective Intelligence In Organizations: A Workshop of the ACM Group*, pages 103–233.

Konijn, E. A., Utz, S., Tanis, M., and Barnes, S. B. (2008). How technology affects human interaction. In Barnes, S. B., editor, *Mediated Interpersonal Communication*, chapter 1, pages 3–13. Routledge New York.

Kou, Y. and Nardi, B. (2013). Regulating anti-social behavior on the Internet: The example of League of Legends. In *Proceedings of iConference*, pages 616–622. iSchools.

Krause, B., Schmitz, C., Hotho, A., and Stumme, G. (2008). The anti-social tagger: detecting spam in social bookmarking systems. In *Proceedings of the 4th international workshop on Adversarial information retrieval on the web*, pages 61–68. ACM.

Kunz, W. and Rittel, H. W. (1970). *Issues as elements of information systems*, volume 131. Institute of Urban and Regional Development, University of California Berkeley, California.

Lange, C., Bojars, U., Groza, T., Breslin, J. G., and Handschuh, S. (2008). Expressing Argumentative Discussions in Social Media Sites. In *Social Data on the Web (SDoW), Workshop at the 7th International Semantic Web Conference*, pages 31–42, Karlsruhe, Germany.

Langlois, G. and Elmer, G. (2013). The research politics of social media platforms. *Culture machine*, 14:1–17.

Lawrence, J., Bex, F., Reed, C., and Snaith, M. (2012). AIFdb: Infrastructure for the Argument Web. In *COMMA*, pages 515–516.

Lee, E.-J. and Shin, S. Y. (2014). When do consumers buy online product reviews? effects of review quality, product type, and reviewer's photo. *Computers in Human Behavior*, 31:356–366.

Leskovec, J. and Faloutsos, C. (2006). Sampling from large graphs. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 631–636. ACM.

Leskovec, J., Kleinberg, J., and Faloutsos, C. (2005). Graphs over time: densification laws, shrinking diameters and possible explanations. In *Proceedings of the 11th ACM SIGKDD international conference on Knowledge discovery in data mining*, pages 177–187. ACM.

Li, G., Uren, V., Motta, E., Shum, S. B., and Domingue, J. (2002). Claimaker: Weaving a semantic web of research papers. In Horrocks, I. and Hendler, J., editors, *The Semantic Web — ISWC 2002*, pages 436–441, Berlin, Heidelberg. Springer Berlin Heidelberg.

Markova, S. and Petkovska-Mirčevska, T. (2013). Social media and supply chain. *Amfiteatru Economic*, 15(33):89–102.

Metzger, M. J. and Flanagin, A. J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics*, 59:210–220.

Moor, A. d. and Aakhus, M. (2006). Argumentation support: From technologies to tools. *Communications of the ACM*, 49(3):93–98.

Munroe, R. (2008). Youtube Audio Preview. URI: http://blog.xkcd.com/2008/10/08/youtube-audio-preview/ (Accessed: 18/07/2014).

Newman, S. and Marshall, C. (1992). Pushing Toulmin too far: Learning from an argument representation scheme. In *Xerox PARC Tech Rpt SSL-92-45*.

O'Keefe, D. J. (1992). *Readings in argumentation*, volume 11, chapter 5, pages 79–91. Walter de Gruyter.

Palau, R. M. and Moens, M.-F. (2009). Argumentation mining: the detection, classification and structure of arguments in text. In *Proceedings of the 12th international conference on artificial intelligence and law*, pages 98–107. ACM.

Panahi, S., Watson, J., and Partridge, H. (2012). Social media and tacit knowledge sharing: Developing a conceptual model. In *International Conference on Information Retrieval and Knowledge Management (ICIKM12)*, pages 1095–1102, Paris, France. World Academy of Science, Engineering and Technology (WASET).

Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society*, 6(2):259–283.

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin UK.

Payne, L. V. and Vaughan, L. (1969). *The Lively Art of Writing*. New American Library.

Peroni, S. and Shotton, D. (2012). FaBiO and CiTO: ontologies for describing bibliographic resources and citations. *Web Semantics: Science, Services and Agents on the World Wide Web*, 17:33–43.

Plato (380BC). *Book V*. The Republic. Basic Books. (Bloom, A.D. Trans. 1991).

Pólos, L. and Hannan, M. T. (2002). Reasoning with partial knowledge. *Sociological methodology*, 32(1):133–181.

Rahwan, I. (2008). Mass argumentation and the semantic web. *Web Semantics: Science, Services and Agents on the World Wide Web*, 6(1):29 – 37. Semantic Web and Web 2.0.

Rahwan, I., Zablith, F., and Reed, C. (2007a). Laying the foundations for a world wide argument web. *Artificial intelligence*, 171(10):897–921.

Rahwan, I., Zablith, F., and Reed, C. (2007b). Towards large scale argumentation support on the semantic web. In *Proceedings of the 22nd national conference on Artificial intelligence*, volume 2, pages 1446–1451.

Reed, C., Wells, S., Devereux, J., and Rowe, G. (2008). AIF+: Dialogue in the Argument Interchange Format. *FRONTIERS IN ARTIFICIAL INTELLIGENCE AND APPLICATIONS*, 172:311.

Resnick, P., Kuwabara, K., Zeckhauser, R., and Friedman, E. (2000). Reputation systems. *Communications of the ACM*, 43(12):45–48.

Rittel, H. W. and Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy sciences*, 4(2):155–169.

Rowe, M., Angeletou, S., and Alani, H. (2011). Predicting discussions on the social semantic web. In *The Semanic Web: Research and Applications*, pages 405–420. Springer.

Ruiz-Iniesta, A. and Corcho, O. (2014). A review of ontologies for describing scholarly and scientific documents. In *4th Workshop on Semantic Publishing (SePublica 2014)*, volume 1155. CEUR.

Schmachtenberg, M., Bizer, C., and Paulheim, H. (2014). Adoption of the linked data best practices in different topical domains. In Mika, P., Tudorache, T., Bernstein, A., Welty, C., Knoblock, C., Vrandečić, D., Groth, P., Noy, N., Janowicz, K., and Goble, C., editors, *The Semantic Web – ISWC 2014*, pages 245–260, Cham. Springer International Publishing.

Schneider, J., Davis, B., and Wyner, A. (2012). Dimensions of argumentation in social media. *Lecture Notes in Computer Science*, 7603:21–25.

Schneider, J., Groza, T., and Passant, A. (2013). A review of argumentation for the Social Semantic Web. *Semantic Web*, 4(2):159–218.

Schneider, J., Villata, S., and Cabrio, E. (2014). Why did they post that argument? Communicative Intentions of Web 2.0 Arguments. In *Arguing on the Web 2.0*, Amsterdam. SINTELNET, European Network for Social Intelligence.

Schopenhauer, A. (1831). *The Art of Always Being Right*. Gibson Square. (Grayling, A.C. Trans. 2004).

Searle, J. R. and Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*, volume 626. Cambridge university press.

Selvin, A. (1999). Supporting collaborative analysis and design with hypertext functionality. *Journal of Digital Information*, 1(4).

Selvin, A., Buckingham Shum, S., Seirhuis, M., Conklin, J., Zimmerman, B., Palus, C., Drath, W., Horth, D., Domingue, J., Motta, E., et al. (2001). Compendium: Making meetings into knowledge events. In *Knowledge Technologies 2001*, Austin TX.

Sherchan, W., Nepal, S., and Paris, C. (2013). A survey of trust in social networks. *ACM Computing Surveys (CSUR)*, 45(4):47.

Shotton, D. (2010). Cito, the citation typing ontology. *Journal of Biomedical Semantics*, 1(1):S6.

Shum, S. B. et al. (2008). Cohere: Towards web 2.0 argumentation. *COMMA*, 8:97–108.

Smith, T. (2009). The social media revolution. *International journal of market research*, 51(4):559–561.

Snaith, M., Bex, F., Lawrence, J., and Reed, C. (2012). Implementing ArguBlogging. In *Proceedings of Computational Models of Argument*, pages 511–512.

Sood, S. O., Churchill, E. F., and Antin, J. (2012). Automatic identification of personal insults on social news sites. *Journal of the American Society for Information Science and Technology*, 63(2):270–285.

Spitzberg, B. H. and Hoobler, G. (2002). Cyberstalking and the technologies of interpersonal terrorism. *New Media & Society*, 4(1):71–92.

Suler, J. R. and Phillips, W. L. (1998). The bad boys of cyberspace: Deviant behavior in a multimedia chat community. *CyberPsychology & Behavior*, 1(3):275–294.

Sundar, S. S. (2000). Multimedia effects on processing and perception of online news: A study of picture, audio, and video downloads. *Journalism & Mass Communication Quarterly*, 77(3):480–499.

Suzuki, S. (2011). Perceptions of the qualities of written arguments by Japanese students. *Written Communication*, 28(4):380–402.

Tindale, C. W. (2007). *Fallacies and argument appraisal*. Cambridge University Press.

Torroni, P., Prandini, M., Ramilli, M., Leite, J., and Martins, J. (2010). Arguments against the troll (position paper). In *1st Workshop on Argumentation in Artificial Intelligence and Philosophy*.

Toulmin, S. E. (1958). *The Uses of Argument*. University Press, Cambridge.

Van Eemeren, F. H. and Grootendorst, R. (2004). *A systematic theory of argumentation: The pragma-dialectical approach*, volume 14. Cambridge University Press.

Van Mieghem, P. (2011). Human psychology of common appraisal: The reddit score. *IEEE Transactions on Multimedia*, 13(6):1404–1406.

Verheij, B. (2005). Evaluating arguments based on Toulmins scheme. *Argumentation*, 19(3):347–371.

Walton, D. (1996). *Argumentation Schemes for Presumptive Reasoning*. Routledge.

Walton, D. N. (1987). The ad hominem argument as an informal fallacy. *Argumentation*, 1(3):317–331.

Wigmore, J. H. (1913). *The Principles of Judicial Proof As Given by Logic, Psychology, and General Experience and Illustrated in Judicial Trials*. Little, Brown and Co., Boston, MA, 1st edition.

Willard, N. E. (2007). *Cyberbullying and cyberthreats: Responding to the challenge of online social aggression, threats, and distress*. Research Press.

Yee, N. and Bailenson, J. (2007). The Proteus Effect: The Effect of Transformed Self-Representation on Behavior. *Human communication research*, 33(3):271–290.

Zanbaka, C., Goolkasian, P., and Hodges, L. (2006). Can a virtual cat persuade you?: The role of gender and realism in speaker persuasiveness. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1153–1162. ACM.

Appendix A

Ethical Approval of Experiments

A.1 Expert Review

This is the ethics submission for the expert review of the ASWO carried out in Chapter 4, and was granted with code 17924.

A.1.1 Ethics Application

A.1.1.1 Pre-Study

Characterise the proposed participants Participants will be academic experts of postgraduate level or above in the fields of argumentation research, ontological modelling, web science or e-bile. Participants will all be public experts with strong profiles in the research community, from the University of Southampton and other institutions. There will be no personal relationship between the primary investigator and the proposed participants; however the primary investigator is a member of some shared communities (through publications and participation in events), such as Computational Models of Natural Argument and Hypertext.

Describe how participants will be approached Participants will be invited via targeted emails:

Dear <participant name>,

My name is Tom Blount and I'm a PhD Student at the University of Southampton, studying argumentation on the social web, how it's modelled, and how perception of arguments affect the course of debate. Currently, I'm conducting an expert review of a new model for describing social argumentation (ethics reference ERGO/FPSE/17924). As an expert in the field of <argumentation research/ontological modelling/web science/e-bile>, I would very much like to ask your opinion of this model in a short (45 minutes) interview over Skype. This will consist of two small exercises to familiarise you with the model, and a series of questions to evaluate the

model itself. Please reply by email if you would be able to help with this study. Alternatively, if you cannot commit to a Skype interview, then I could send you the materials to look at in your own time.

Thank you in advance for your help,
Tom

Tom Blount (PhD candidate)
Web and Internet Science
Electronics and Computer Science
University of Southampton, UK

A follow up email containing additional information (including consent forms) will be sent to any participants who indicate they are interested in participating:

Dear <participant name>,

Thank you for expressing your interest in taking part in this study.

During the study you will be asked to read an explanation of current argument modelling techniques (5-10 minutes) and then complete a short "talk aloud" exercise implementing some of these techniques (5-10 minutes). You will then be asked to read a short explanation of proposed extensions to this model (5-10 minutes) and to repeat the "talk aloud" exercise using these extensions (5-10 minutes). You will then be asked ten questions covering the completeness, clarity and coherence of these extensions (5-10 minutes). Please read the attached Participant Information document for further information on what is required and how the data will be managed. If you are happy to participate, please fill out the attached Consent Form and return it via email, along with a list of times you would find most convenient to participate. You are also welcome to browse and familiarise yourself with the attached Questionnaire in advance. If you would like to participate, but cannot commit to a Skype interview, you may go through the Questionnaire in your own time and return the results via email. If you have any questions, please don't hesitate to email me. Your participation in this study is voluntary and you may withdraw at any time without your legal rights being affected.

Thank you in advance for your help,
Tom

Tom Blount (PhD candidate)
Web and Internet Science
Electronics and Computer Science
University of Southampton, UK

Describe how inclusion and/or exclusion criteria will be applied (if any) Participants must be of postgraduate level or above in the fields of argumentation research, ontological modelling, web science or e-bile

Describe how participants will decide whether to take part Participants will be provided the Participant Information and Consent Form documents along with the invitational email to ensure they make an informed decision on whether or not to take part. They can then reply to the email to elect to take part.

A.1.1.2 During the Study

Describe the study procedures as they will be experienced by the participant Participants will be asked to read an explanation of current argument modelling techniques (5-10 minutes) and then complete a short "talk aloud" exercise implementing some of these techniques (5-10 minutes). Participants will then be asked to read a short explanation of proposed extensions to this model (5-10 minutes) and to repeat the "talk aloud" exercise using these extensions (5-10 minutes). Participants will then be asked ten questions covering the completeness, clarity and coherence of these extensions (5-10 minutes).

Identify how, when, where, and what kind of data will be recorded (not just the formal research data, but including all other study data such as e-mail addresses and signed consent forms) An interview of the participants completing the "talk-aloud" exercises and answering the study questions will be recorded (as will any questions relating to their understanding of the tasks).

Participant questionnaire (Reproduced in Appendix B)

A.1.1.3 Post-Study

Identify how, when, and where data will be stored, processed, and destroyed The data is relevant to the study purposes because it focuses solely on evaluating the theory and execution behind the extended argumentation model. The data is adequate because it covers the principals involved, and evaluates the completeness, clarity and coherence of the model and the data is not excessive because it covers only those aspects required. The data will be processed fairly because the participants will have given explicit consent. The data's accuracy is ensured because the interview will be recorded verbatim. Data will be stored on the Investigator's desktop. The data will be held in accordance with University policy on data retention. Data files will be protected by password and stored on investigator's desktop; desktop will be protected by password and stored in a secure/card-accessed research lab. The data will be destroyed by secure deletion at the conclusion of the study: 31st December 2015. The data will be processed in accordance with the rights of the participants because they will have the right to access, correct, and/or withdraw their data at any time and for any reason. Participants will be able to exercise their rights

by contacting the investigator (e-mail: tb12g09@soton.ac.uk) or the project supervisor (e-mail: dem@soton.ac.uk). The data will contain personal data in that participants could be identified by the recording of their voice. The data will have no other associated personal information (name/email/etc.). Consent forms will be linked to the data by an identification number No data will be transferred outside the European Economic Area (EEA).

A.1.2 Participant Information

Study Title: Expert Review of Argumentation Model

Researcher: Tom Blount

Ethics number: 17924

Please read this information carefully before deciding to take part in this research. If you are happy to participate you will be asked to sign a consent form.

What is the research about? This research aims to evaluate an extended means of modelling social argumentation on the web.

Why have I been chosen? You have been chosen as an expert in the field of argumentation research, ontological modelling, web science or e-bile

What will happen to me if I take part? You will be asked to read an explanation of current argument modelling techniques (5-10 minutes) and then complete a short "talk aloud" exercise implementing some of these techniques (5-10 minutes). You will then be asked to read a short explanation of proposed extensions to this model (5-10 minutes) and to repeat the "talk aloud" exercise using these extensions (5-10 minutes). You will then be asked ten questions covering the completeness, clarity and coherence of these extensions (5-10 minutes).

Are there any benefits in my taking part? It is hoped that this research will add to current knowledge of online argumentation, and may aid the improvement of online discussion

Are there any risks involved? There are no risks associated with your participation

Will my participation be confidential? Because the review will take the form of a recorded interview, it is not possible to anonymise your data. However, data will be stored on a password protected computer in a secure lab. You will retain the right to access and correct your data, and your right to request removal of their data; to do this, contact the primary investigator, Tom Blount (tb12g09@soton.ac.uk) or project supervisor David Millard (dem@soton.ac.uk).

What happens if I change my mind? Your participation is voluntary and you may withdraw at any time without your legal rights being affected.

What happens if something goes wrong? Should you have any concern or complaint, contact the primary investigator, Tom Blount, if possible (tb12g09@soton.ac.uk); otherwise, please contact Dr Martina Prude, Head of Research Governance (02380 595058, mad4@soton.ac.uk).

A.2 Perception of Argument Study

This is the ethics submission for the expert review of the ASWO carried out in Chapter 6, and was granted with code 18150.

A.2.1 Ethics Application

A.2.1.1 Pre-Study

Characterise the proposed participants Participants will be members of the general public, 18 years of age or above, and familiar with social media.

Describe how participants will be approached A link to a website with information about the study will be publicised via social media sites, to attract users of social media sites. They will then be able to read this information and express their consent to the study if they are willing.

Describe how inclusion and/or exclusion criteria will be applied (if any) Participants must be 18 years of age or above, and be familiar with social media

Describe how participants will decide whether to take part Participants will be provided the Participant Information via a website; once they understand the study, they may make an informed decision on whether or not to take part by filling in the Consent Form, also via the website, before moving to the online survey.

A.2.1.2 During the Study

Describe the study procedures as they will be experienced by the participant Participants will be asked to read a series of posts taken from the public social media profiles of five news operations. For each post, participants will be shown up to two replies to the post, and asked to evaluate these posts on several attributes (see Questionnaire attachment) using Likert scales.

Identify how, when, where, and what kind of data will be recorded (not just the formal research data, but including all other study data such as e-mail addresses and signed consent forms) No personal data about participants will be collected or stored; the annotations they make to the presented social media posts will remain anonymous.

Participant questionnaire (To be answered on a Likert scale)

1. This post is coherent/easy to understand
2. This post contains credible information
3. This post makes a persuasive argument
4. This post is entertaining
5. This post is offensive
6. How likely would you be to reply to this post?
7. How likely would you be share this post (to friends/followers/etc.)?
8. How likely would you be to up-/down-vote this post?
9. How likely would you be to report this post (or its author) to moderators?

A.2.1.3 Post-Study

Identify how, when, and where data will be stored, processed, and destroyed

Data will be collected and stored on a secure university virtual-machine for the duration of the study. Data will be transferred to and processed on the investigators desktop, which will be protected by password and stored in a secure, card-accessed research lab. The data will be held in accordance with University policy on data retention. The data will be processed fairly because the participants will have given explicit informed consent. No data will be transferred outside the UK or European Economic Area (EEA). Participants will not be able to withdraw or access their data as it will not be personally identifiable. The anonymised data will be retained to further other research as specified by the funding body (EPSRC).

A.2.2 Participant Information

Study Title: Perception of Social Media Comments

Researcher: Tom Blount

Ethics Number: 18150

What is the research about? This research aims to establish how people perceive different types of argumentation, from different areas of the social web

Why have I been chosen? You have responded to a call for participation, are familiar with social media, and are aged 18 or over

What will happen to me if I take part? You will be asked to read a series of posts taken from the public social media accounts of five news operations. For each post, you will be shown up to two replies, and asked to answer a series of questions

Are there any benefits in my taking part? It is hoped that this research will add to current knowledge of online argumentation, and may aid in the improvement of online discussion

Are there any risks involved? There are no risks associated with your participation

Will my participation be confidential? No personal data about you will be collected or stored. All answers given will be completely anonymous.

What happens if I change my mind? Your participation is voluntary and you may withdraw at any time without your legal rights being affected. However, you will not be able to withdraw or modify your data (as it will not be possible to identify it).

What happens if I have questions, or something goes wrong? Should you have any question, concern or complaint, contact the primary investigator, Tom Blount (tb12g09@ecs.soton.ac.uk); the project supervisor, David Millard (dem@ecs.soton.ac.uk); or the Head of Research governance (02380 5959058, rginfo@soton.ac.uk)

Appendix B

Expert Information Sheet

The following is a reproduction of the information sheet provided to experts conducting the review in Chapter 4. Note that some of the diagrams used in the information sheet below predate those used to illustrate Chapter 4.

B.1 Proposal

This work aims to extend the current methods for modelling web based argument to take into account additional social features and differentiating between "logical" argument that focuses on (purported) facts and "rhetorical" argument that focuses on influencing the perception of participants in the eyes of the audience. This hopes to make the modelling of "eristic" argument (argument for the sake of argument) more complete, clear and consistent.

B.2 Existing Models

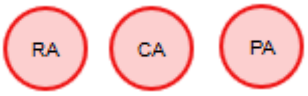
B.2.1 Argument Interchange Format

The Argument Interchange Format (AIF) is a framework for representing argumentation as a directed graph (Chesñevar et al., 2006), modelling information "nodes" and the relationships (such as inference or conflict) between them. In their work on an extension to the AIF, dubbed AIF+, Reed et al. differentiate between these logical relations and the actual words spoken during the debate (Reed et al., 2008). Table B.1 shows an overview of these nodes and how they are used in the AIF(+).

B.2.2 Semantically Interlinked Online Communities

The Semantically Interlinked Online Communities project (SIOC), a semantic-web vocabulary for representation social media, aims to enable the cross-platform, cross-service representation of data from the social web (Breslin et al., 2006). This allows for semantic representations of Sites, which hold Forums, which contain Posts, authored by a UserAccount (explicitly *not* a person, as a person can own and manage more than one UserAccount). Table B.1 shows an overview of the nodes used in SIOC.

Table B.1: Description of nodes in model

Name	Description	Node	
I-node	**Information** nodes represent a (purported) piece of information, data, or claim		AIF
S-nodes (RA-, CA-, PA-nodes)	**Scheme** nodes denote a logical connection between I-nodes, respectively an **inference**, a **conflict**, or a **value preference**		AIF
YA-node	**Illocutionary anchor** nodes tie the information and logical structure of an argument with the spoken or written locution		AIF+
L-node	**Locution** nodes represent the actual words that are spoken or written by participants		AIF+
TA-node	**Transition** nodes represent links between locutions. **Note:** this is adapted by the ASWO to denote transitions between locutions that do not add information nodes, but still further the debate (such as prompting for more details, evidence, etc.)		AIF+
U-node	**User-account** nodes denote the account the user uses to contribute		SIOC

B.2.3 Examples

B.2.3.1 Syllogism

A syllogism is an example of reasoning in which two premises are used to draw a conclusion. Figure B.1 shows a syllogism of the form *"All men are mortal. Socrates is a man. Therefore Socrates is mortal"*.

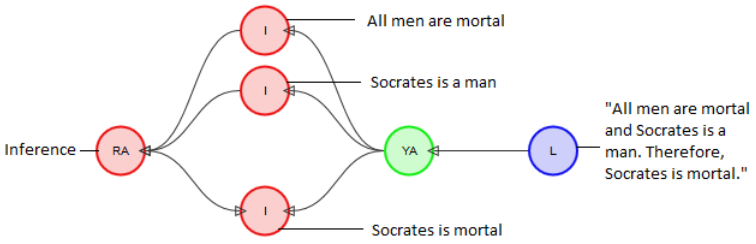


Figure B.1: Example of a syllogism: *"All men are mortal. Socrates is a man. Therefore Socrates is mortal"*

B.2.4 Exercise 1

Please read the following sample arguments and describe (aloud, if you are being interviewed face-to-face) how you would model them using the AIF(+) and SIOC. You may find sketching them on a piece of paper useful. If you are feel unsure of how to model all or part of one of these samples, move on to the next part.

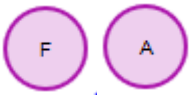
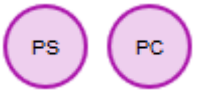
1.
   - **User 1:** *The tech industry is often biased against women*
   - **User 2:** *@User1 You would say that, you're a woman*
   - **User 3:** *@User1 **** off and die you ****ing nazi before I come and **** you up*
2.
   - **User 1:** *Guns killed 33,000 people last year, they need to be banned*
   - **User 2:** *@User1 And a lot of those were minors*
   - **User 3:** *@User2 According to who?*
3.
   - **User 1:** *What does Barack Obama call illegal aliens? Undocumented democrats!*
   - **User 2:** *@User1 You're so stupid you probably went to the library to find Facebook*

B.3 Argumentation on the Social Web Ontology

The principal features from the AIF and SIOC ontologies are combined alongside the means to model rhetorical tactics in the Argumentation on the Social Web Ontology (ASWO). The

principal focus here is the inclusion of the social impact of arguments made and the use of rhetorical support and attack (Blount et al., 2014, 2015a). Table **??** shows an overview of the additional nodes used to model social impact.

Table B.2: Description of nodes added to the model

Name	Description	Node
P-node	**Persona** nodes denote a person's social "character" that they assume during a discussion	
F- and A-nodes	**Faction** and **Audience** nodes represent groups of personas	
PS-, PC-nodes	**Personal Support** and **Personal Conflict** nodes support/attack personas or groups rather than pieces of information	
Im-node	**Implication** nodes imply a relationship that may or may not exist. Can be combined with a PS- or PC-node to denote positive or negative implication	

ASWO

B.3.1 Examples

B.3.1.1 *Ad hominem*

Ad hominem ("to the man") arguments attack a person's character, without attacking their argument. However, they can be a viable tactic in rhetorical debate and can introduce both new I-, CA- and PC-nodes to the structure when modelled.

Figure B.2 shows a "reasonable" *ad hominem* argument (Walton, 1987), such as *"You don't have any qualifications in that area, don't make such broad statements"*

Figure B.3 shows a more aggressive tactic that disparages someone's argument and them as a person, such as *"They're an idiot, don't listen to them"*

Figure B.4 shows an abusive argument that contains no information, instead attacking the person directly and trying to shut them out of the debate, for example *"**** off and die!"*

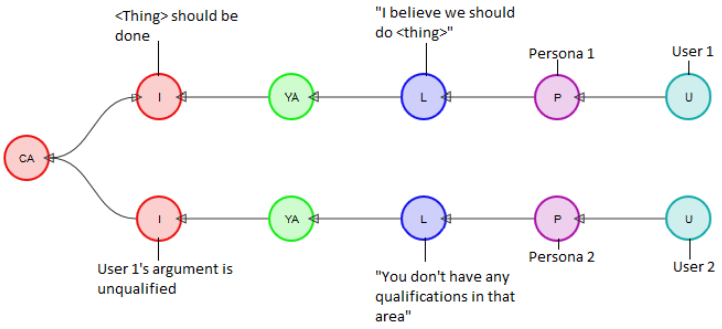


Figure B.2: Example of a reasonable *ad hominem* attack

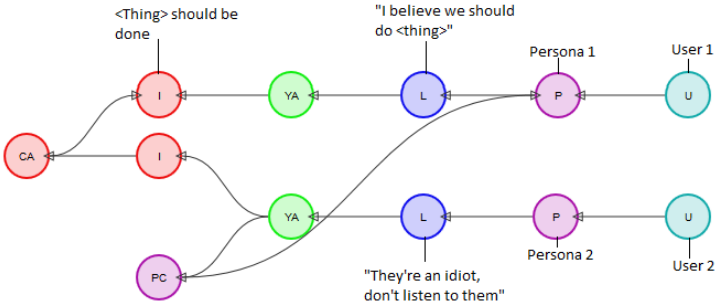


Figure B.3: Example of an *ad hominem* attacking both persona and argument

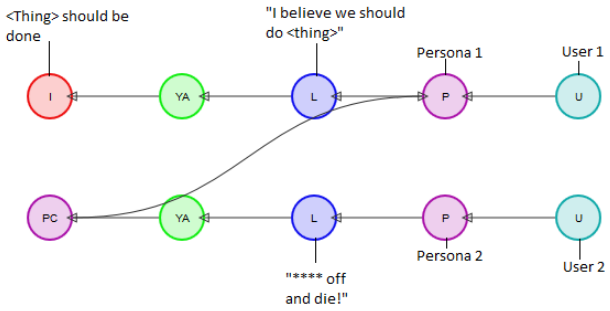


Figure B.4: Example of an abusive *ad hominem* attack

B.3.1.2 Appeal to Consensus

Appeal to consensus is the fallacy that because a claim is popular or widely-held, it is true. An example of this can be shown in Figure B.5.

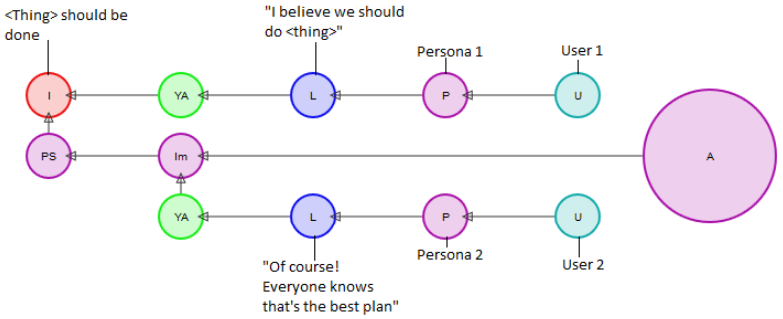


Figure B.5: Example of Appeal to Consensus

B.3.1.3 Association Fallacy

The association fallacy is the notion that because a person is associated with, or shares the views of, an undesirable group, their claims are wrong. An example of this can be shown in Figure B.6.

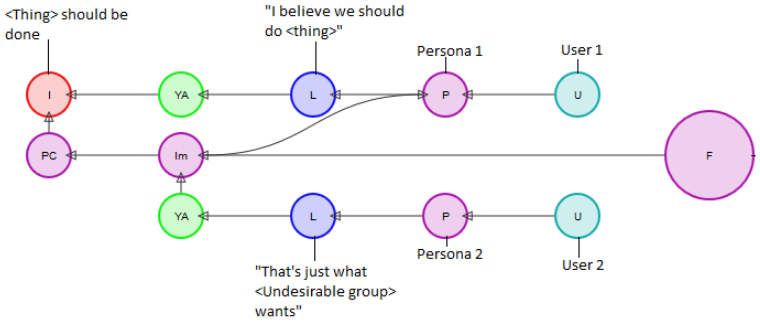


Figure B.6: Example of the association fallacy

B.3.1.4 Appeal to Humour

Appeal to humour is a technique by which a participant in the debate attempts to ingratiate themselves with their audience by making a joke about the situation as shown in B.7. This can be coupled with an *ad hominem* attack when the joke is made at someone else's expense.

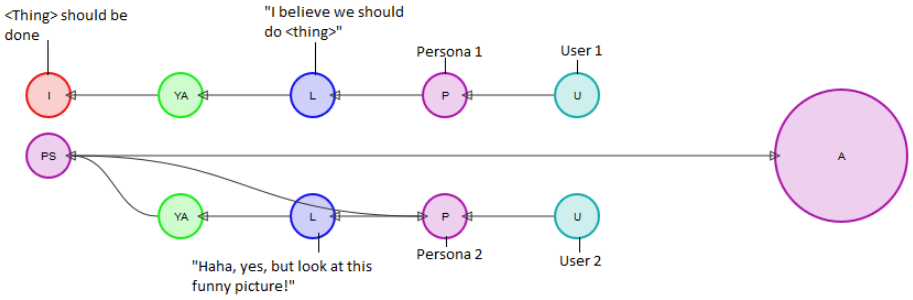


Figure B.7: Example of Appeal to Humour in the model

B.3.2 Exercise 2

Please read the following sample arguments and describe (aloud, if you are being interviewed face-to-face) how you would model them using the additional nodes added by the AWSO. You may find sketching them on a piece of paper useful. If you are feel unsure of how to model all or part of one of these samples, move on to the next part.

1.
   - **User 1:** *The tech industry is often biased against women*
   - **User 2:** *@User1 You would say that, you're a woman*
   - **User 3:** *@User1 **** off and die you ****ing nazi before I come and **** you up*
2.
   - **User 1:** *Guns killed 33,000 people last year, they need to be banned*
   - **User 2:** *@User1 And a lot of those were minors*
   - **User 3:** *@User2 According to who?*
3.
   - **User 1:** *What does Barack Obama call illegal aliens? Undocumented democrats!*
   - **User 2:** *@User1 You're so stupid you probably went to the library to find Facebook*

B.4 Questions

1. Why do you feel social argumentation is, or is not, important to model?
2. What, in your opinion, are the challenges of modelling social argument?
3. Are threatening and/or abusive comments something that should be considered social argumentation? If not, where should the line be drawn?
4. If yes, how do you feel these threatening and/or abusive comments should be included?
5. To what extent did the ASWO capture different elements of argumentation? What do you feel is missing?
6. Were there parts of the ASWO you felt were unclear? In what way?
7. Do you feel the ASWO is consistent with the AIF?
8. Do you feel the ASWO is internally consistent?
9. If two people were to separately model the same argument using the ASWO, do you think they would achieve the same result? Do you feel this is important?
10. Do you have any other comments about the implementation of this model?

Appendix C

Social Media Post Feature Distributions

C.1 Number of Comments

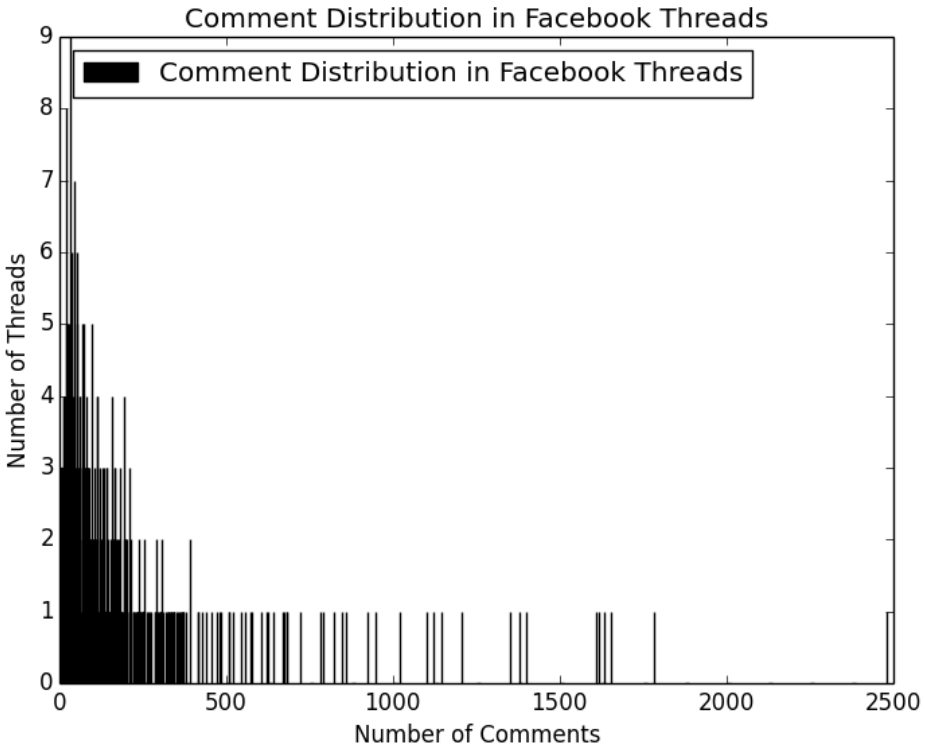


Figure C.1: Number of comments per thread on Facebook

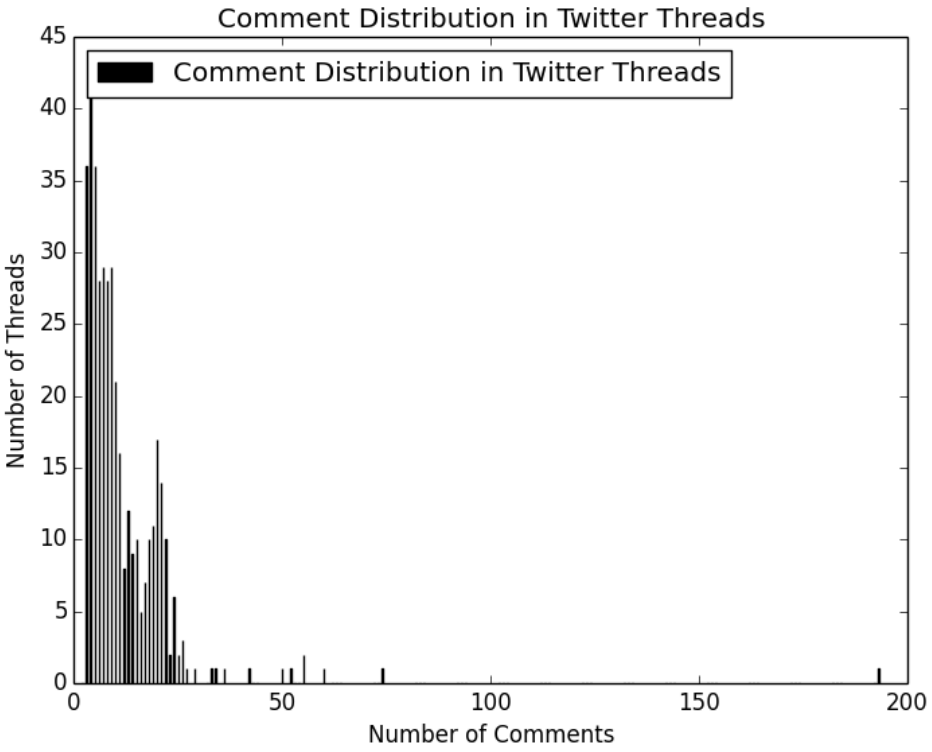


Figure C.2: Number of comments per thread on Twitter

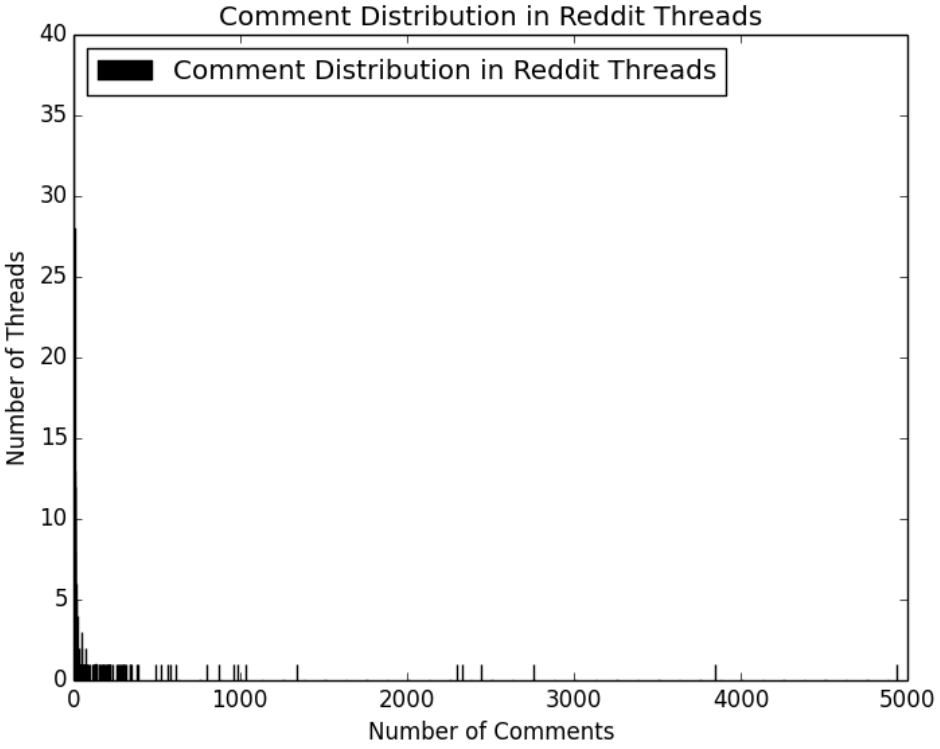


Figure C.3: Number of comments per thread on Reddit

C.2 Length of Comments

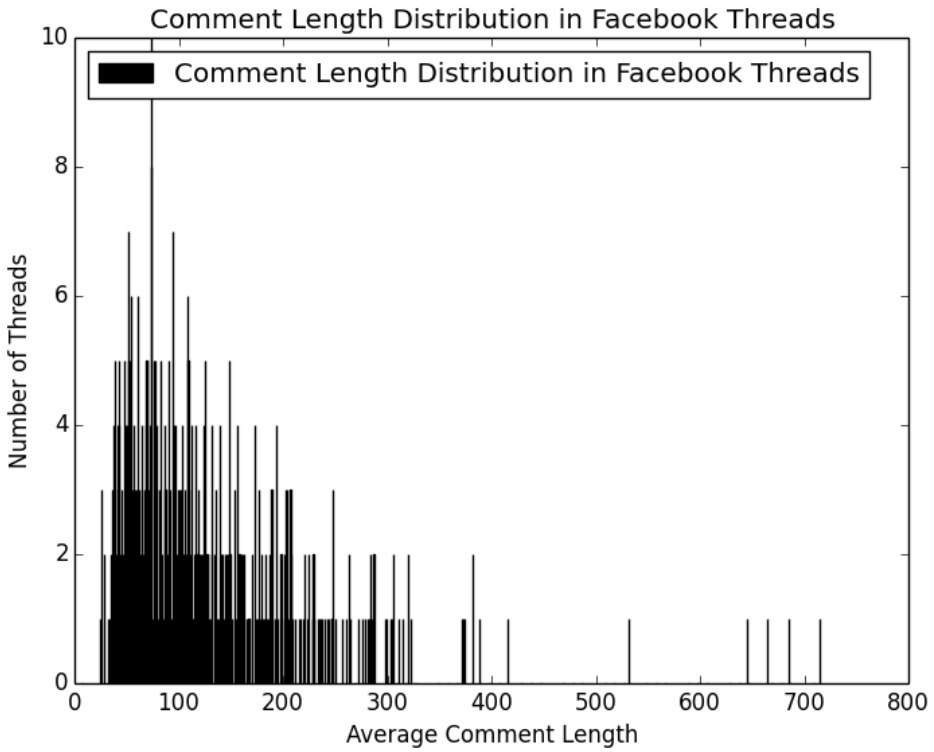


Figure C.4: Average length of comments per thread on Facebook

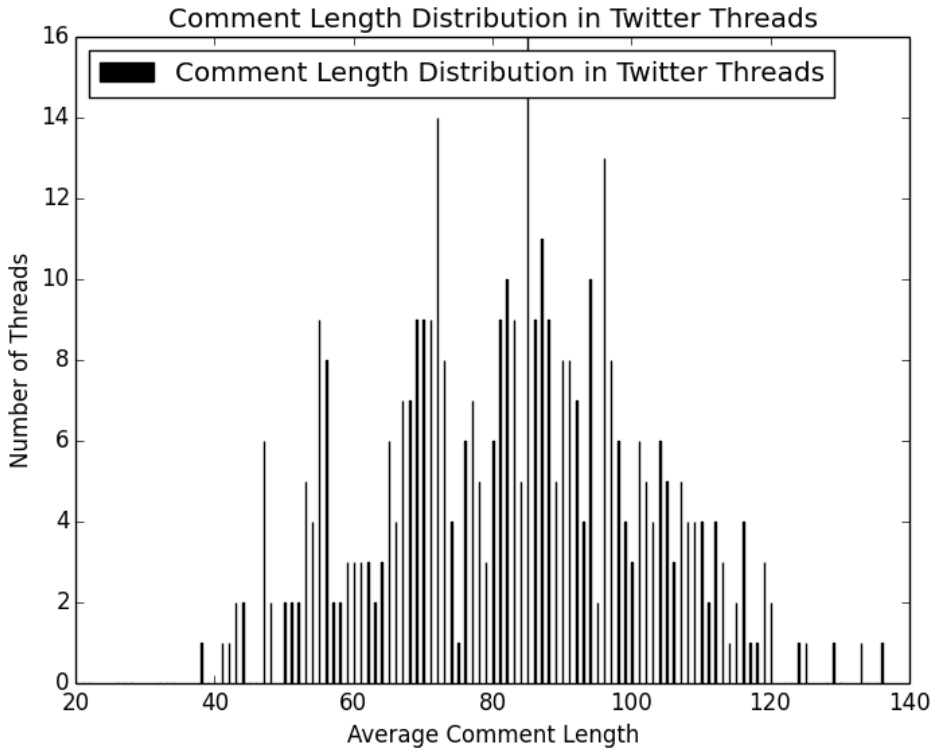


Figure C.5: Average length of comments per thread on Twitter

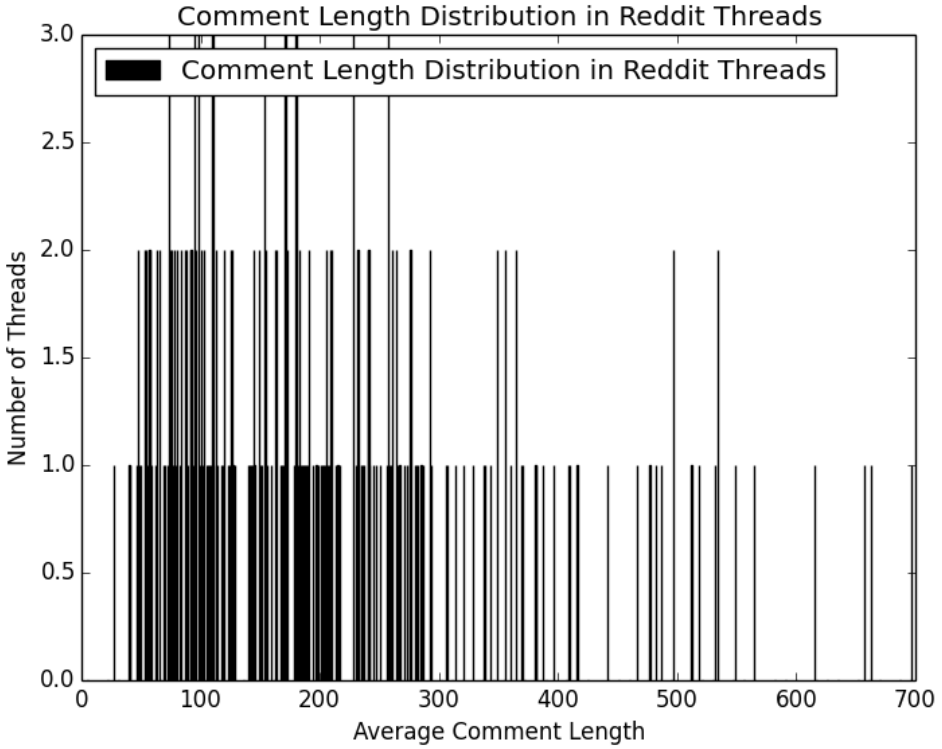


Figure C.6: Average length of comments per thread on Reddit

C.3 Comments per User

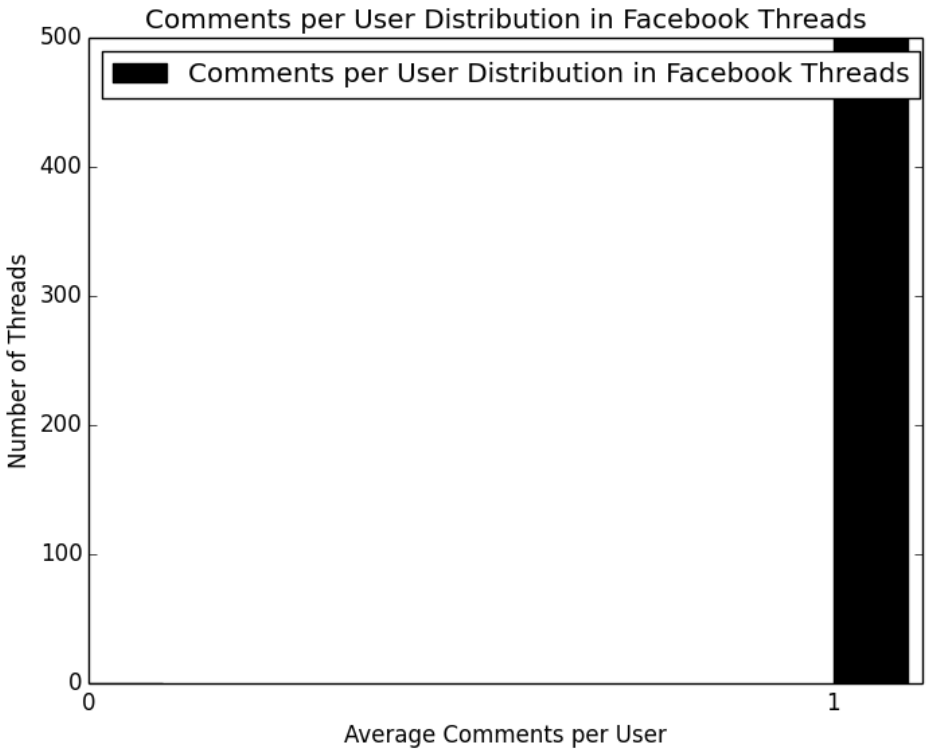


Figure C.7: Average comments per user per thread on Facebook

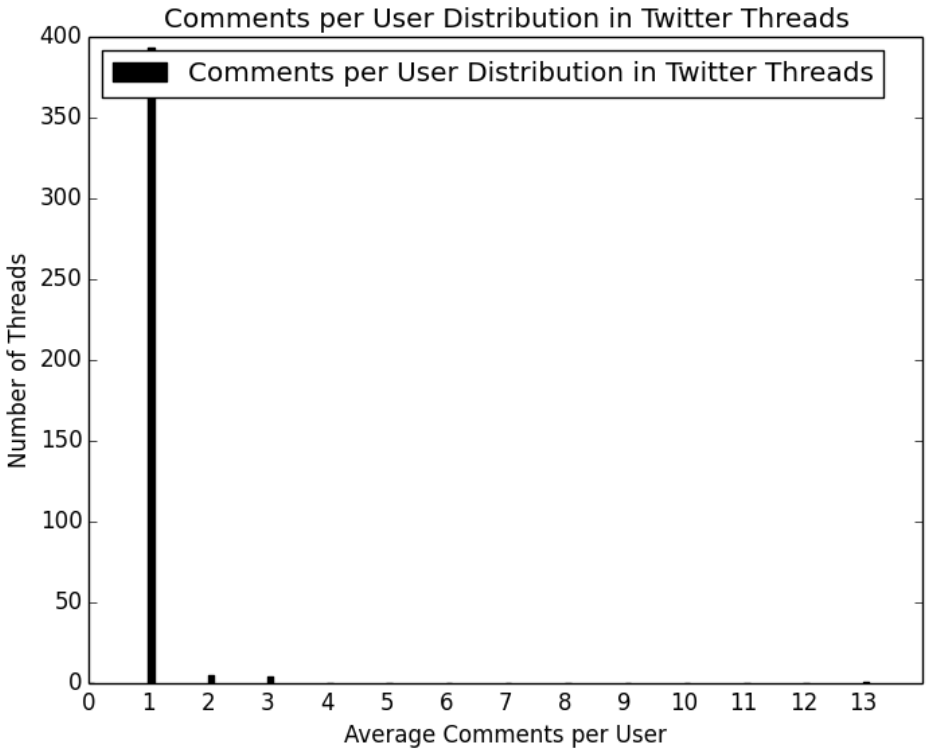


Figure C.8: Average comments per user per thread on Twitter

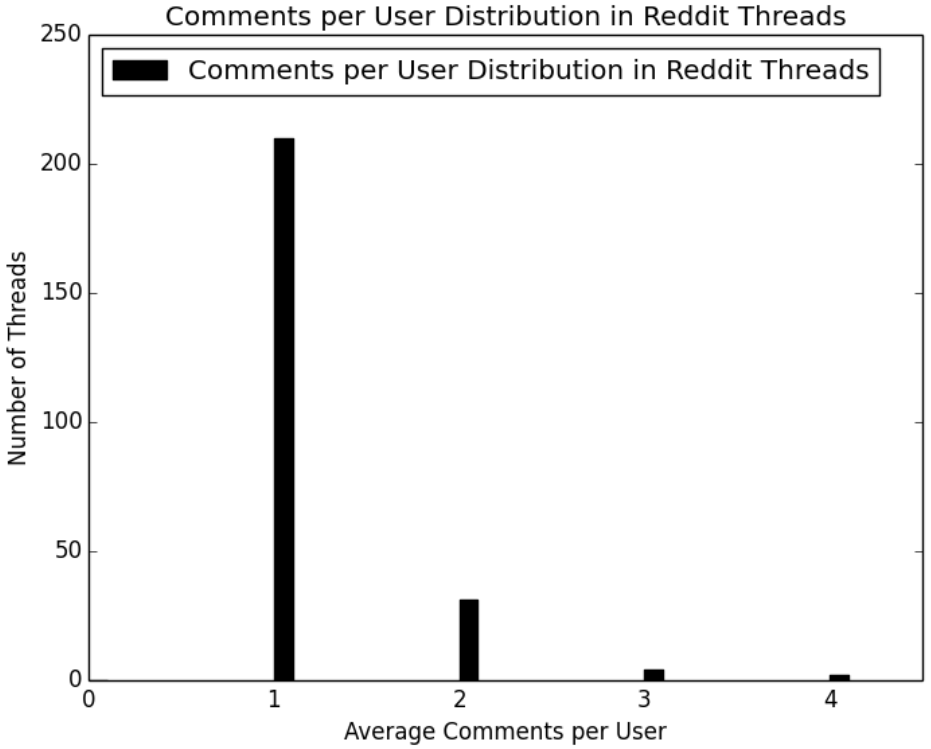


Figure C.9: Average comments per user per thread on Reddit

C.4 Replies within Thread

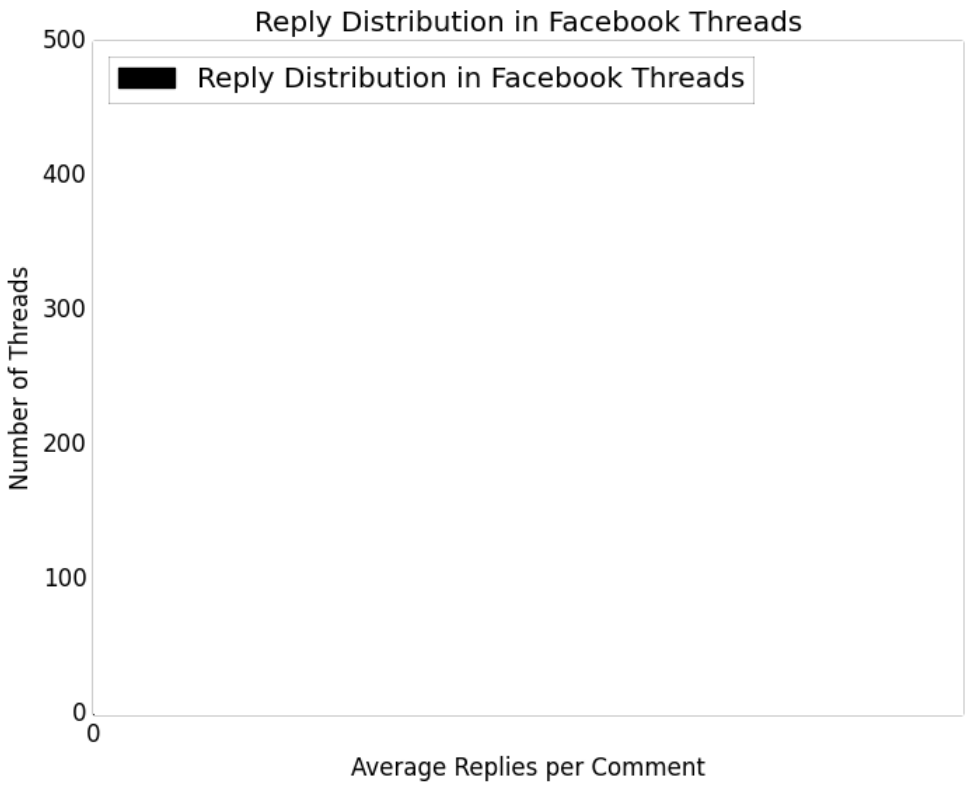


Figure C.10: Internal replies per thread on Facebook

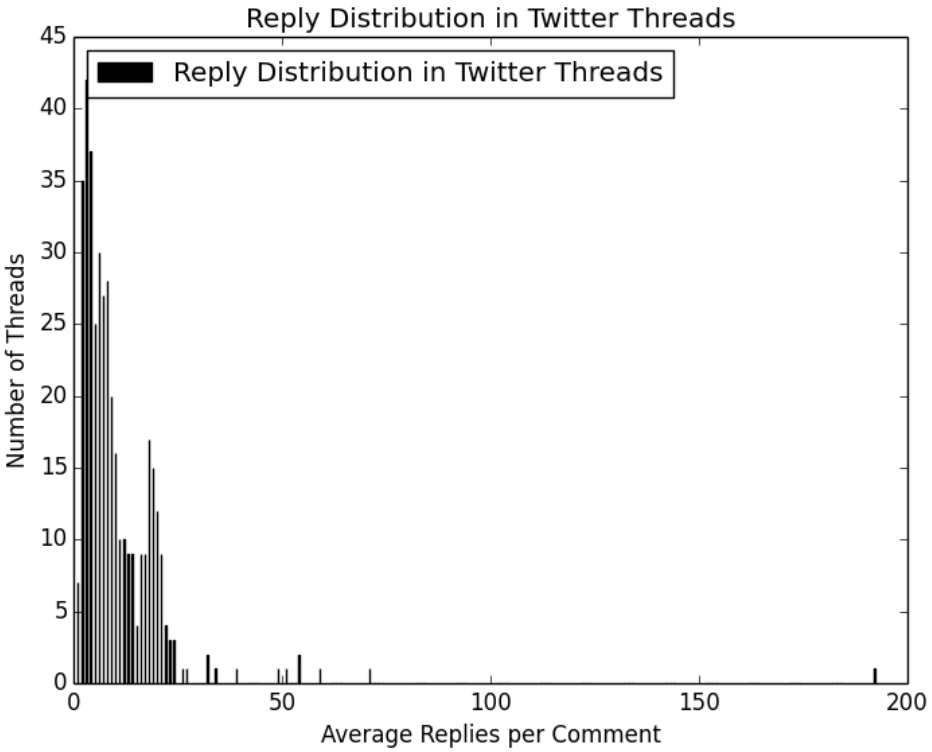


Figure C.11: Internal replies per thread on Twitter

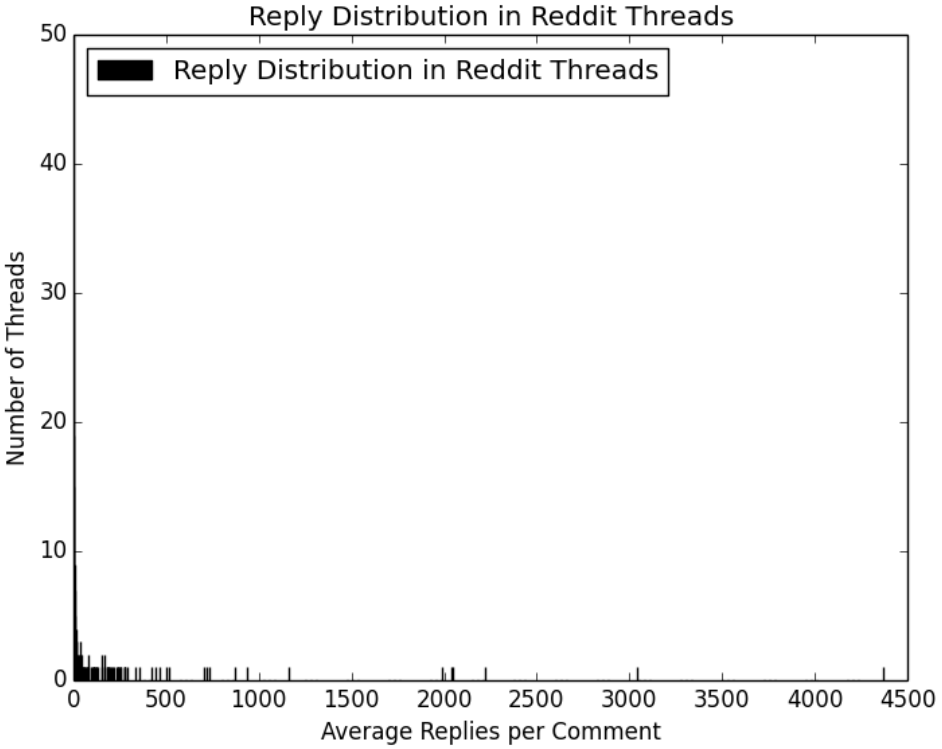


Figure C.12: Internal replies per thread on Reddit

Appendix D

Sample of Social Media Threads

D.1 Facebook

05:12 24/02/16

US election 2016: Trump wins Nevada – Strengthening his position in the Republican presidential race.

More: http://bbc.in/1Qbfb0J #NVCaucus

05:13 24/02/16

Crap

05:13 24/02/16

Oooops

05:13 24/02/16

No sevens on this win. Nothing but a bunch of blanks and no payoff. Stupid is as stupid does for those votes, but Cruz and the others aren't any good either.

05:14 24/02/16

The reason I will not vote for TRUMP is because he is too much like ME and I know how dangerous I would be if I was President. Trump appeals to everything that I am, that is Bad . When I hear him speaking for most of what he says its like hes speaking whats in my head and when i am angry there is a lot going on in my head that if I acted on would hurt a lot of people.

I'm trying to Change that part of me and I am trying to hold back those emotional responses. If Trump is President he will release and encourage every bad thing in me. He is like the figurative devil on my shoulders, imagine voting for and electing that little small voice.

Donald Trump is the embodiment of every negative trait and character flaw that I and his supporters have. He is a freaking mirror of all our Character flaws and everything that's bad about us that needs to Change. Unfortunate for America Donald Trump and Ted CRUZ are going to be the next Administrative leaders Ted might very well become the figurative good Angel on our shoulder. Can you imagine what that would look like. I can see it and it actually scares me.

Trump speaks to every negative trait in me. He appeals to my anger and everything in me that's negative. If he becomes president he will represent every weak flaw in me that I am trying so , so very hard to change. Voting for Trump would be the most irresponsible thing I could ever do.

Just because Donald says what most of us who like him think, does not justify voting for him. The only good thing about Trump is that he is not a Politician. He is not apart of the good Ol' boys club. Other than that there is not too much left. My Educated opinion is Trump will only make the situation worse, as would anyone who makes decisions based on Anger. Do we need a leader like that? Will that really make America Great Again?

Trump by his folly will only strengthen and empower an enemy that already is blind by there own fury and there own anger there own leaders who speak to the worst parts of them. An enemy that is blinded to how they are being used and manipulated by there own emotional manipulators. When will it ever stop. We need a leader that can answer that Question and actually Implement a plan that can bring peace to everyone involved.

Donald Trump is going to be the next President and I guarantee that within 6 months of him being in office EVERYONE that voted for him will sorrowfully regret having done so. Only then will you realize how irresponsible it is to respond in anger to this very messed up situation. And by then it will be too freaking late.

I might be wrong about Ted Cruz....

05:15 24/02/16

Acho que vou come ar a acreditar nas testemunhas de jeov , o fim do mundo aproxima−se!

05:15 24/02/16

Interesting times we live in. I'd really like to see Trump and Sanders win their respective party's nomination. I think they might both force our two political parties to reform and produce better candidates. It'd be nice to see them shake up DC too, they're both political outsiders and it seems most establishment politicians aren't very fond of either of them.

05:15 24/02/16

Say 'NO' to Trump Sheeples

05:16 24/02/16

Get ready world for American refugees to begin flooding Europe and the rest of the world. That's if we will even be allowed to leave with his giant wall and all.

05:17 24/02/16

Really Trump on his way to having his finger on nukes . . . not funny.

05:17 24/02/16

Yes to Trump is no to the NWO!

05:19 24/02/16

Does he ever button the bottom of his suit coats or sport coats ?? Can he ??

05:19 24/02/16

Americans proved They're baap of all racist

05:19 24/02/16

Right on!!

05:19 24/02/16

Racist chalk up another win.

05:19 24/02/16

king of liar

05:20 24/02/16

Trump is not a Republican, just like Bernie is not a Democrat.

05:20 24/02/16

Urghhhh... Another trumping post

05:21 24/02/16

'Trump the dumb' is the kind of person i lose control over my sanity whenever i come across. Humanity is being trampled upon where Trump the Dumb is.....

05:24 24/02/16

Please spare a thought for the rest of us, America....It's like the world's crappiest comic just scored 7 days with the world's crappiest agent, and with six more pub gigs to go, he's already searching around for the steering wheel that drives Airforce one....(good news... the heat's shifted away from a relieved Putin, the world's second crappiest comic)

05:24 24/02/16

https://www.facebook.com/inthenowrt/videos/578195305664176

05:25 24/02/16

lol, it means haters hate him only on FB.

05:25 24/02/16

Trump is winning evangelicals, moderates, northeasterners, southerners, westerners and his rival democrats are still attacking each other. Barring something completely unforeseeable, Trump will be the next nominee lol

---

05:27 24/02/16

America is going to hell in a gasoline soaked hand basket!

05:30 24/02/16

Buba Vanga predicted that Obama will be the last president of the US after him its going to fall, so Trump will become president and he will run the US down into the ground.

05:30 24/02/16

All nasty noise and tasteless theater no positive substance but he appeals to the ugly emotions of those who flock to an authoritarian figure who fuels the flames of their hate, fear, and vulgarity.

05:32 24/02/16

America will suffer for another 8 wasted years!!!! What is wrong with people???? He's our GOP Obama only worse!!!

05:33 24/02/16

When I heard Trump was running, I thought it was all a big joke. When I listen to him speak, I think OMG, this man in nuts, then he keeps winning these states, now I think there is lead in all the drinking water in this country because Americans have lost their minds. This will be the fall of America. What fools.

05:33 24/02/16

Stefano Fiore he's slowly on the rise mate

05:34 24/02/16

Republians will fail if they nominate Trump as their candidate.

05:34 24/02/16

What a tool

05:35 24/02/16

trump will win and it proves most of america is uncivilized

05:35 24/02/16

online mobile bussiness
work from home
zero investment
100% legal bussiness
100% pure income
work on fb & whatsapp only
work only 30 minuts in a day
instant withdrawal in your account
earn monthly 15,000 to 20,000
for this bussiness you need only smartphone & Net

digital india .....
.....

...
.....

....
.
.

.
Join WhatsApp no. 9088069580 ł send
...
.
..

05:36 24/02/16

http://www.nationalreview.com/article/431694/donald-trump-global-bully

---

05:38 24/02/16

Never trust a political aspirant whose tie tip reaches his dick.

05:39 24/02/16

Awful news

05:39 24/02/16

Venessa Thor

05:40 24/02/16

https://www.facebook.com/czech.WeAreHereAtHome/videos/1532678907033072/?pnref=story

---

05:42 24/02/16

USA is going down.....

05:43 24/02/16

I'm ashamed to live is this country....

05:46 24/02/16

Soon jolly ol England will have to reclaim these colonies

05:47 24/02/16

I HAVE UNFOLLOWWED CNN COZ THEY POST 24*7 about DONALD DUCK TRUMPSHIT

05:47 24/02/16

You just ran a article on BBC 24, around 15 minutes ago, regarding Apple fighting the FBI over the court ruling to unblock the iPhone of the killer Syed Rizwan Farook. The article makes a terribly gaffe in it. It shows a clip of Mark Zuckerburg speaking on this subject, but a caption came up underneath stating' Donald Trump – Facebook CEO'. It then immediately shows a clip of Donald Trump speaking on this subject, correctly captioned as 'Donald Trump – Republican Presidential Candidate'. It is bad enough that Mr Trump is having success in his bid, as I consider some of his comments at times to be ugly, unwanted and crass. But to now having top news agencies make glaring blunders in editing their articles and mistakenly give Mr Trump credit for being CEO of Facebook, really takes the biscuit!

05:48 24/02/16

Am just following from a distance!

05:49 24/02/16

Are US people going to follow India??? Electing Trump is same as P.M MOODI an extremist.....hats off another extremist

---

05:50 24/02/16

OMG I can't believe it

05:52 24/02/16

Noooooooo. We are better than this.....

05:57 24/02/16

@Tommy please can we swap country...?

05:57 24/02/16

Mein trump, an American neo–Nazi's rise through opposition and hardship.

06:00 24/02/16

The Democrats are cross fingers to hope Teump being the head of Republican .. so they will win the final presidancial elecrion in November. .

06:02 24/02/16

Be vewy vewy afwaid.

06:05 24/02/16

trump is nothing but the right choice,the messiah to make us great again,he will make Americans proud to be americans

---

06:07 24/02/16

America z responsible state trump z not eligible for president be aware American .

06:08 24/02/16

More and more, my fellow Americans prove how incredibly STUPID they truly are.

06:10 24/02/16

I hope trump wins so America will finally get the final war it so craves and burn

---

06:12 24/02/16

http://youtu.be/1iDRu2q2cRw

06:13 24/02/16

Hahahha pipo actually voted for this man wow..intresting

06:16 24/02/16

TRUMP for President ! (Better than the muslim traitor and his terrorist friends CAIR that they have suffered for so long)

06:17 24/02/16

He d candidate 2 beat at the final if anyone can. Americans need such rugged guy now to help build that great nation again whkich all of us all over d world depended greatly on for provision and protection. Go on Trump you're d best.

06:18 24/02/16

https://www.youtube.com/watch?v=V3miuaOWsj8

06:20 24/02/16

gud 2 go.

06:22 24/02/16

> What started as joke is fast becoming a reality and it so crazy to think that this man may become the POTUS in no distance time. You all should not be surprise when he came out after winning and said he ran just for heck of it and that he has no idear of what to do now as the prisident.

06:23 24/02/16

> Shaking my head, that's about all I can do when I read about his victories. I seriously question the intelligence and mindset of his supporters. God help us.

06:24 24/02/16

> what they lost there dame mind right

06:25 24/02/16

> Go to hell

06:28 24/02/16

> If I have to choose between Donald or Hilary I think I'll just go back to my country Haiti still don't think that safe

---

06:28 24/02/16

> America they will go down if he win

06:34 24/02/16

> Trump 4 real

06:37 24/02/16

> His money goes to killing wildlife: https://www.facebook.com/photo.php?fbid=874280182685478&set=a.347612158685619.82769.1000

06:40 24/02/16

> I am.

06:42 24/02/16

> Hahaha God help you if he win the election

06:48 24/02/16

> After Hitler, after Ben Laden, this is Trump from KKK – for a world of peace don't vote for him ! USA is not Trump

---

07:00 24/02/16

> Really America ?

07:13 24/02/16

Iam supporting Trump because his comedy remarks, he said, "i will build a wall in mexican border and mexico will pay it

07:26 24/02/16

Woke up to make some Tea and what great news! What makes it so funny is reading all of the irrational, panicked, misinformed, exaggerated and moronic comments in this thread.

07:33 24/02/16

This win also reflects the true character of Americans.. If they can support him its coz they believe in his opinions and actions. No wonder they are killing each other daily

07:37 24/02/16

Just by reading all these comments, one can understand why Americans are voting for Trump!!

07:40 24/02/16

Nooooo

07:40 24/02/16

if him win Clinton don't have to run against him.
that no mek sense. when the election every body just go vote.

07:41 24/02/16

Sanaya Shikari WHAT IS THISS

07:44 24/02/16

Most unusual to have someone running for a political position, speaking what a lot of people are scared to say, for fear of prosecution. Well in the UK they would. Be prosecuted, locked away and forgotten about. I was joking Dragon. I was. lol (TmA)

07:46 24/02/16

oh dear. This man will ruin America.

07:47 24/02/16

OK let's who is best, Adolf Hitler or Donald Trump... ??

07:49 24/02/16

Whoever votes for him deserve what they get!

07:49 24/02/16

and it funny how the media a dig up things and a press things on ms Clinton and them not doing trump nuttin. every minute. email..now speach fi money.Bengazi no one can stop a terrorist attack them can only try to prevent it. she do her best.. she is a strong woman. stronger than the iron lady. keep up the good work ms clinton punch them left jabs and right and one uppercut in the end. we from the island of Jamaica love u.

07:57 24/02/16

trump hates muslims n almost all american christians n jews hate muslims so its a simple calculation he is the next president of USA.

08:07 24/02/16

Hater of human kind ? Some says ?

08:08 24/02/16

He must be paying them to vote, the man is a total joke to humanity.

08:11 24/02/16

#trumpcheats
#trumpdirtytricks
Ted Cruz Unites, leads, apologizes for minor faults, accepts responsible. A true conservative leader. Trumps doesn't accept responsibility, divides groups, and almost never apologizes. Nothing to add only subtracts. Trumps obviously has many Non–Leader liabilities.

... & This country is going to hell in a golden woven basket. Lol Hey it's not over to the foul mouthed boy sings. But until that day it's better to laugh than cry.
For Trump's insults the USA needs a good bar of American made soap. "Guard the door of your lips." So Trump said that he could "murder" someone? My supporters are loyal"(Paraphrasing) I like comedy, but why even joke about violence?And he should apologize for calling THE GREAT POPE FRANCIS "DISGRACEFUL", that would be nice. And bring more respect for himself, Roman Catholics & the Great USA.

CRUZ 2016 Conservative fairness & unity. America First. Security & Constitution.

Tired of Trumps Tantrums & Tactics?? Donald Trump is an angry divisive man with some money in an wannabe Statesman's empty suit. He is not a Unifier in these United States of America, he's a divider.
Petty politics AA ball. Trump better put on his baseball cap again, the big leagues are next.

Here's a defense & idea, the press/media is a double edge sword. What's old is new, new old. Very easily you can find current & archived negative press stories about Trump the last 50 yrs. there must be plenty plenty of NY press articles about Trump divorces, evictions, play boying, politician games, casino shenanigans, lawsuits, at least one thousand, 1000 Trump Wrongdoings, lies, and just not nice behavior. Have a team find them, highlight, word search the word "lie","stole","lawsuits" briefly in each of the one thousand articles found. Lol. Unleash it on block all at once on your website & everywhere, the press. And prove that he(Trump) is the liar, wrong doer, not nice person. And archived anti Trump 1000 Article Trump media FACTUAL review. I'd do it, just an idea.

CRUZ 2016 Leadership.
Accountability, Resolve. More Action, Less Talk.
Ted Cruz 2016 A Proven Advocate for Freedom. Getting results. Policy Results & discourse. Understsnding the party not attacking it. Bringing the Conservative Republican Party together, not dividing it.
–Trump: No track record. Proven Insults. –

Americans don't want a strong leader, they want a good one.
Ted Cruz 2016 Ł Building Bridges Before Walls. Trust Ted you won't be fired! Trust Ted, he won't fire you. Ted Terrific.

For CRUZ 2016 & supporters::
–"Never, Never, Never give up."
Winston Churchill

For Trump & supporters:
"Courage is what it takes to stand up and speak; courage is also what it takes to sit down and listen."–
Winston Churchill

08:16 24/02/16

that's something been added to their food so they have lost reason WAKE UP AMERICA before its to late

---

08:17 24/02/16

is it only the rest of the world that can see him for who he is? have they put something in the US water supply?

08:26 24/02/16

Putin on one side, Trump potentially (but hopefuuly not) on the other. Each with their nuclear buttons. At least Kennedy and Kruschev were both sane

---

08:26 24/02/16

i am really confuse now, people voting for trump to be a president or wants trump clown to keep running a circus show for them and entertain them. but which one?

08:29 24/02/16

Trump wins again!
The positive thing about him is he isn't just ambitious. He wants to tackle America's problems head on, so you have to admire him for that. But he seems to have negative about whole ethnic groups. You can't do that. He should be targeting the criminal's and drugs people in all USA communities. His ideas seem not to be thought threw and not workable or fair. But he certainly isn't boring!

08:32 24/02/16

Help

08:33 24/02/16

What the heck is wrong with them?

08:46 24/02/16

The best choice for the US

08:46 24/02/16

Nice short

09:04 24/02/16

Express your opinion using Ripplear dot com.

09:11 24/02/16

What a troll.

09:14 24/02/16

you must be so sad BBC. And 46% from the hispanic votes? How your media propaganda trying to make trump look like he hates mexicans has failed. We get our news from truthful sources now, your time is over

09:16 24/02/16

what i get to realised is that trump keeps loosing on social media but winning in d real world in U.S... just saying my view frm Africa

09:23 24/02/16

Never mind young Americans this is dumb Americans

09:34 24/02/16

"It is the absolute right of the State to supervise the formation of public opinion."
Joseph Goebbels

09:38 24/02/16

Jesse Boeve Dino Kadric Bertram van der Aa Bram Verbaas NOBODY CAN STOP US! WE WILL MAKE AMERICA GREAT AGAIN!!!!

09:48 24/02/16

love you trump .you can only save America and this planet .vote for trump

---

09:51 24/02/16

Congrats .keep it up .I support u

09:53 24/02/16

Top job Trump! How'd you pull that one off? God?

10:01 24/02/16

I feel sick.

10:17 24/02/16

GOP must be shaking by now I believe...Go Trump#2016

10:31 24/02/16

They will get what they deserve, god help us all

10:40 24/02/16

Oh whats going on, it must be something wrong

10:51 24/02/16

I don't get it, if he's winning it means the mass public of these states are agreeing with the racist, facist and extremely vile statements Trump continuously makes? These people are simply government controlled patriots systematically taught to believe that everything there government does has no exterior motive other than 'God bless America' lol what a joke

---

10:55 24/02/16

Go Trump... Go

11:05 24/02/16

QUESTION: Why does the media report about John Kasich sticking his foot in his mouth about women in the kitchen, but the media does not report the FACT that Kasich is a proven COWARD regarding his very own daughters and would be DANGEROUS as Commander in Chief?
Refer to the web sites http://NotKasich.com/ and http://ABCsOfBetrayal.com/ for the facts as reported by the Columbus Dispatch including audio recordings of those who got caught during that 10-year Conspiracy of Criminal Negligence, Deception, Fraud, and Cover-Ups committed against thousands of consumers involving millions of dollars.

11:32 24/02/16

Trump wins, more guns, more wars, more racism, killings, more intolerance , just bloody great! Exactly what the world needs right now aye !!?

11:40 24/02/16

Those litle countries in America its not relevant, two bigest states in America gona decide who is going to win presidental elections those states California and Texas gona decide i dont even know why people overeacted so much about this...

11:46 24/02/16

His appearance says enough... People like him wiped out the American Indians.. Calling them savages! To grab land that belonged to the native Indians! Barbaric!

12:01 24/02/16

Very very scary let's all pray the US comes to its senses but with 95% of the US population only ever seeing their own state you can sort of forgive them for not understanding how scared the rest of the world is right now

---

12:24 24/02/16

let him also rule change in presidenny can also change som thing as in develpment.

12:29 24/02/16

Oh my good grief, there is no hope for humanity!!! Hitler was also allowed to go to run a Country and look what he did!!!

12:46 24/02/16

Naden Scarfone this frightens me

12:54 24/02/16

TRUMP WILL CAUSE. INTERNATIONAL CONFLARATION IF HE SUCCEEDS ON THE PATH HE IS THREADING . HE SHOULD BE THROWN OUT BY DEMOCRATIC AMERICANS !

13:07 24/02/16

Dont forget we had Thatcher in this country.

13:59 24/02/16

Nasir Babangida

14:07 24/02/16

Yes trump love it .from the UK

14:16 24/02/16

Ibrahim Abdel–Hafiz

14:20 24/02/16

If Trump wins then its true America has stable supply of fools!..i know that can't happen.

14:21 24/02/16

American version of Mar Roxas.

14:23 24/02/16

OHM America where are you going! That is terrible.

14:33 24/02/16

Money bailed him out; hope he burns out.Trumped up tit.

14:52 24/02/16

God Bless Trump, from us here in the UK! Ł

14:57 24/02/16

Syria ,Iran ,Turkey ,Uganda ,zimbawe ,Eq–Guinea and Gambia.Jast play so hard that Trump should not become the president of USA.He will bulldoze u 4rom right ,left and center.This man iz crazy.

15:39 24/02/16

Good decision ,thanks Americans as you look forward to usher in a new democratic changes because the foreign policies have to be worked on and Trump is more ideal

15:49 24/02/16

Can't believe voters are stoking the ego of one of the maddest people on Earth.

16:06 24/02/16

16:13 24/02/16

Good luck Trump, from supporters in Wales, UK.

16:53 24/02/16

I hate this world. Wake up sheep, they're leading you to the slaughter!

17:06 24/02/16

on beharth of working ppl Bristol UK we support you Donald J. Trump

17:09 24/02/16

America must of lost the plot even more than normal god fekin help us if this thing gets in charge

17:18 24/02/16

Trump is being called "the next great messiah who will make this country great again." Funny. That's what they said of Adolph Hitler.

17:54 24/02/16

I find this more scary than isis

17:59 24/02/16

good luck

18:06 24/02/16

Dear rest of the world: pray for us. God help us if he wins the presidency.

18:18 24/02/16

Trump all the way! Ł

18:32 24/02/16

Que bien vamos Donald trump.. tienes que ganar en todo usa

18:33 24/02/16

Gooo Donald..

18:59 24/02/16

Will it fit on the mantlepiece?

19:32 24/02/16

Please nooooo

19:46 24/02/16

Go Donald....

20:16 24/02/16

He's a coward and i hate when people spread rumours about muslims that not only causes hatred among other nations but also spreads unfriendly relations which is known to all those nations that had extreme fights between just because of rumour the whole territory is destroyed it not only kills human beings but it destroys everything that once was achieved...

20:31 24/02/16

MERHABA ARKADALAR, PARAYA HT YACI OLAN DOSTLARIMIZIN KES NL KLE OKUMASI GEREKEN B R YAZIDIR.
renci dostlar m z, memur dostlar m z, paraya ihtiyac olan aile sahibi dostlar m z, k saca paraya ihtiya duyan herkesin bana ula mas n temenni ederim.
Tek yapman z gereken bana ula makt r. Profilime mesaj atabilirsiniz. Takip etmekten bir ey kaybetmezsiniz.
yi vakitler dilerim arkada lar.

21:14 24/02/16

They deserve help,research your subject first ,he could save USA ,Main Street media is not honest sorry

---

21:18 24/02/16

Ramin Mahmoodi Jeez

21:29 24/02/16

Scarey!!

21:52 24/02/16

I can't believe that the American people are so backwards at coming forwards that they are even considering letting this moron in the White House

---

22:09 24/02/16

Erin Geiger

22:14 24/02/16

The people in the United States. .had enough. ..and people are voting in big numbers for. The first time. ..they are winning. ..Trump is helping them. ..

22:20 24/02/16

so there is dis majority who think like dis man,

22:22 24/02/16

Well looks like im never coming to vegas again.

22:57 24/02/16

PURE TRADITIONAL HERBS WHICH ADHERE TO THE INTERNATIONAL NORMS
Powerful Spiritual Traditional Herbalist Healer
Experienced in Ancestral healing and spell casting, Astrologers, African Medicines, Ritualism, Herbalist healer, Spiritual healer, Native healer, Philosophy, Traditional healer.
Phone No: +2349057218254
Email : Greatsalemtemple@gmail.com
Facebook ID: GREAT SALEM TEMPLE

Herbal remedies, holistic healing;
1. Sandawana oil for all spells.
2. Bring Back Lost Lover, even if lost for a long time
3. Remove Bad spells from homes, business &customer attraction etc.
4. Get Promotion you have desired for a long time at work or in your career.
5. Remove the Black spot that keeps on taking your money away
6. Find out why you are not Progressing in life and the solution
7. Eliminate in Family Fights among est each other
8. Ensure excellent school grades even for children with Mental Disabilities
9. Stop your Marriage or Relationship from breaking apart
10. We destroy and can send back the Tokoloshe if requested
11. We heal Barrenness in women and disturbing menstruation
12. Get you marriage to the Lover of your choice
13. Loose and Gain weight
14. Guarantee you win Family Protection Spells.
15.Sangoma & Pastors who need more Powers.
Are your family members getting sick mysteriously?
Are your family members not making any progress in life with their career, in school or with their business?
Protection Spells to protect and keep your family safe and help them to be successful in life and to prosper

22:58 24/02/16

Be fun see him leading the un haha

23:08 24/02/16

His been bank rap 2 time he will do it with the USA

23:14 24/02/16

I'm all aboard the Trump 2016 train!!

23:17 24/02/16

ARMAGEDDON a bad feeling about this.

23:58 24/02/16

This is joke .

00:32 25/02/16

This is what happens when you put drugs in a nations waters to drug everyone and make them robots, no longer can people think , but then again, you didn't have a better competition did you?

00:35 25/02/16

trump is a fool.

01:13 25/02/16

digital india ł Andriod mobile RS 15000−50000/ . ł Advertising . ł ł ł ł ... ł whatsApp no. (9780383935) ł "JOIN" send .

03:13 25/02/16

To all the sane people of Nevada, you have my sympathies.

03:52 25/02/16

The Holy Bible never say in anywhere that #Jesus is #God,
But the #Bible prove that #Jesus is the son of #God and was sent by #God the #Father as #Prophet on #Earth to save the whole world.

The devil was haven #Authority on Earth before #Jesus #Matt 4:1 to 11 but after his #Crucifixion and #Resurrection all #Authority in #Heaven and on #Earth has been given to him by #God the #Father #Matt 28:18
Now #Jesus is #Lord in #Heaven and a #King on the whole #Earth

While #God the #Holy #Father is #King of #Kings and #Lord of #Lords

Now the devil as a ruler now on Earth has no Authority on us but can only deceive us to do Bad. The #Bible never lie just that we are all deceive by the Roman Empire with the Teachings ...

Stay Awake b'cus the Kingdom of God is @Hand

04:27 25/02/16

Trump would like to take us back to pre–Magna Carta days...

05:37 25/02/16

Hmmm, welcome to reality. If African leaders will invest in Africa. All these will not matter. Have we heard of Japan, S. Korea Singapore complained about his harsh statements.

08:30 25/02/16

join the illuminati to be rich and famous in the whole world, whasapp +2348143559477 for help now

---

09:15 25/02/16

TUCK FRUMP

09:17 25/02/16

What Ever Trump say .... he Really does not Mean Anything .... if it Does .... then its not practical ....In today's world and Emerging Global Market. .....He is a business Man .... he is selling his Each and every word. .....& people Are buying it .... Nearly 3 hundred thousand American Muslims peace is jeopardise ... and is going to windows up A racial discrimination ...from one End to other End ....America is going to start A war Or A civil war Inside On Who is **Native American ** who has A right to live ...His latest word we will going to fill lot of bad dude in Jails ...Oh Mr Cow Boy ... they Are Already filled will Afro American s.....?

09:26 25/02/16

He is a good person which be president of .u.s.a

10:06 25/02/16

Donald Trump is a joke, just like the US election system

10:06 25/02/16

Once

10:06 25/02/16

Once trump speaks against Muslims he's going to win

D.2 Twitter

18:36 19/02/16

Kesha 'in tears' after judge denies her release from Sony where producer allegedly raped her https://t.co/thYQgcH9AW https://t.co/XBnz5x5fIJ

18:37 19/02/16

> @dailymailceleb #sonysupportsrape

20:12 19/02/16

> @Btwsts @DailyMailCeleb you do realise thats slander without proof, If sony wanted too they could sue you. People have been sued for less.

20:33 19/02/16

> @Btwsts what rape?

19:34 19/02/16

> RT @DailyMailCeleb: Kesha 'in tears' after judge denies her release from Sony where producer allegedly raped her https://t.co/thYQgcH9AW ht

---

19:36 19/02/16

> @DailyMailCeleb @MailOnline E chi ?

19:38 19/02/16

> She should have know the "casting couch" deal clincher is still practiced in some pop-media orgs.
>
> @DailyMailCeleb @MailOnline

19:40 19/02/16

> @DailyMailCeleb @MailOnline no restraining order?

20:33 19/02/16

> @DailyMailCeleb @MailOnline so she was raped but did not file charges, bull shit.

20:46 19/02/16

> @DailyMailCeleb @MailOnline see, judges can be stupid, too. No wonder it took so long for Cosby accusers to come forward. That sucks, Kesha

---

D.3 Reddit

15:24 21/02/16

Brexit against Scotland's wishes would 'almost certainly' trigger independence referendum, warns Nicola Sturgeon

---

15:27 21/02/16

> You could trigger an independence referendum if Cameron said he doesn't much care for the flavor of Irn Bru

---

19:41 21/02/16

> Irn Bru....
>
> ..It's pig fuckingly good.

00:39 22/02/16

> that moment of deep disappointment when you realize it wasn't a can of lucozade

---

15:52 21/02/16

> 1 they had a referendum and decided to stay, you can't just referendum whenever you feel like it
> 2 Scotland would not gain straight entry to the Euro, there are a few other members who didn't want them with the oil money so without it they would have a battle to get in
> 3 Nicola sturgeon is a cunt

16:49 21/02/16

> You actually can call for a referendum whenever you want.

17:07 21/02/16

> Like France, Ireland etc.

18:52 21/02/16

> 1) yes, yes you can. Whether that referendum actually leads to anything is another story. In saying that, if we voted for independence and then weren't allowed to by Westminster there would be civil war in some shape or form.
>
> 2) Scotland would have a much easier time getting in than Ukraine, for instance. We already follow all of the policies and regulations to the letter.
>
> 3) She is, but she's *our* cunt and I trust her/the SNP a hell of a lot more than the other cunts playing political panto in London.

21:17 21/02/16

> Easier than Ukraine but how long did that take
>
> The thing I don't get is the hate sturgeon kicks up about the English
> 99% of English or Welsh could not give a fuck if your Scottish or not, working in Scotland during the lead up to the vote she got it all very us against them!
> If you don't like Cameron then say that but don't try to turn one country on another

21:28 21/02/16

> It's got minute amount to do with minute amount of English people and a massive amount to do with the English establishment.

21:49 21/02/16

> It's the uk establishment not English!
> If you work it out on votes per representative in Westminster snp has a very good representation, a lot better than labour conservative lib dem or ukip

21:57 21/02/16

> Except for the fact that we have a government that only 46% of population of the UK voted for. Great representation there!

22:07 21/02/16

> Democracy. Not saying I voted conservative or even like the toff nosed prick but there were over 25 different choices nationwide plus independents

> With that many options 40% is a big margin
>
> In America it's a 2 horse race and you can have a leader only 51% votes for

---

22:28 21/02/16

> *FPTP, not the
>
> most representative of democratic systems. Not really that democratic at all once you start thinking about it.

11:07 22/02/16

> Enjoy being governed by Berlin.
>
> I will laugh.

19:30 21/02/16

> The fuck you say, bruv?

11:06 22/02/16

> true
>
> true
>
> I lol'd but true nevertheless.

16:01 21/02/16

> This is the best tl;dr I could make, [original](http://www.independent.co.uk/news/uk/politics/eu-referendum-brexit-scottish- reduced by 69%. (I'm a bot)
>
> *****
>
> > The Scottish First Minister and SNP leader said if England voted to leave the EU while Scotland voted to remain there would be an "Inescapable" shift in public opinion towards independence to guarantee the country's continued EU membership.
>
> > "If, a couple of years later, we find ourselves, having voted to stay in the EU, being taken out against our will, I think there will be many people – including people who voted No in 2014 – who would say the only way to guarantee our EU membership is to be independent."
>
> > Ms Sturgeon acknowledged that an independent Scotland in the EU would have to negotiate its border arrangements if the remainder of the UK was outside the EU. She added that said she has "No proposals" to use her new devolved powers to top up benefits for migrants, but said EU migrants have had a positive impact on the UK economy.

```
*****
[**Extended
    Summary**](http://np.reddit.com/r/autotldr/comments/46vtfa/brexit_against_scotlands_wishes_wou
    |
    [FAQ](http://np.reddit.com/r/autotldr/comments/31b9fm/faq_autotldr_bot/
    "Version 1.6, ~38302 tl;drs so far.") |
    [Theory](http://np.reddit.com/r/autotldr/comments/31bfht/theory_autotldr_concept/)
    | [Feedback](http://np.reddit.com/message/compose?to=%23autotldr "PMs
    and comment replies are read by the bot admin, constructive feedback is
    welcome.") | *Top* *keywords*: **vote**^#1 **EU**^#2 **Scotland**^#3
    **referendum**^#4 **campaign**^#5
```

19:37 21/02/16

```
tl;dr but "brexit" has got to stop. I can't hear that word every day until
    June. Please.
```

22:52 21/02/16

```
Yes please, this is worse than Brangelina.
```

19:57 21/02/16

```
The oil is a bonus.
```

20:08 21/02/16

```
Can we also vote to kick Scotland out? Oil is the only thing you have and
    that's worth dick right now, so leave all you want.
```

00:47 22/02/16

```
well they've got bagpipes and tartan...oh wait they got those
    from ireland...um trump golf courses?
```

20:36 21/02/16

```
How quickly can a referendum be called anyways?
```

21:54 21/02/16

```
that arsey dwarf is (one of ) the reasons why we've got a tory government
```

23:00 21/02/16

```
If every single person in Scotland voted labour, the Tories still
    would have won
```

00:25 22/02/16

> the snp claimed labour would need them to govern, and that scotland would have the sway over a future labour government.
>
> great for her home audience, but the tories used it to make labour look like their puppet– a british parliament working for scotland.
>
> so scotland gets all those mps but as you say, not enough to decide any policies AND they ensured a tory victory.
>
> edited for my shocking spelling

22:03 21/02/16

> This is getting beyond a joke with this Independence crap now, The SNP and more specifically Sturgeon are looking for just about any excuse to force through another Independence vote.
>
> The majority voted no and were vindicated in that vote when it was shown later that the SNPs predictions for North Sea oil were out massively (based on oil being $100+ per barrel) The SNP predictions for 2016/17 Oil revenue was 8 billion, That number is now as low as 500 million.
>
> http://www.prospectmagazine.co.uk/blogs/george–magnus/the–snps–economic–case–is–draining–aw
>
> http://www.ft.com/cms/s/0/ccac5894–a337–11e5–a035–96e9dfdf9fff.html
>
> http://www.telegraph.co.uk/news/uknews/scottish–independence/12052565/Independent–Scotland–
>
> It's got so bad with the SNP incessant talk of another independence vote David Cameron could be seen eating a Tunnock's Teacake and Sturgeon would demand an independence vote, Maybe she should spend more of her time focused on the important politics for Scotland and not this ridiculous idiotic blind drive towards an independence that the majority of Scots have already said no to.
>
> Edit: Here come the downvotes by the hardcore independence mentalists who can never see any bad in what the SNP do as long they get their Freeeeeddoooooommmm.

01:21 22/02/16

> To pretend that Brexit wouldn't effect Scotland is ridiculous. To believe that the leader of the SNP wouldn't mention Independence, over concerns for what Brexit would do to Scotland, is also ridiculous.
>
> Besides, the Independence vote was very, very close, and you can't force us to be your friends forever.

02:04 22/02/16

> it was a third of the eligible electorate, two thirds did not vote for independence (44% my arse)

05:09 22/02/16

> At these oil prices, an independent Scotland would be very poor.

22:19 21/02/16

> Not British, but why would the UK go for Brexit, knowing that this is the case?

01:32 22/02/16

> Hurry up and leave then FFS. Tired of her constantly threatening to leave every time she disagrees with something.
>
> At this point scotland leaving is inevitable if a referendum can be called every couple of years.

01:41 22/02/16

> Don't worry. When the SNP sinks the Scottish economy, they will turf the SNP out.

11:04 22/02/16

> If you can't build a tram system how can you build a country?

04:43 22/02/16

> they could leave the uk tomorrow and for the next 200 years they'd still blame the english every time they found a vegetable in their dinner

---

06:28 22/02/16

> Let them go. I'm tired of the Scots whining.
>
> There is no utopia South of the border that they are being denied.
>
> Put up a border enforce passport control and let them be governed from Berlin by Merkel.
>
> I'm sure Merkel will look after them.
>
> It's a pity the rest of the UK doesn't get a vote, they would definitely be out.

Appendix E

Perception Experiment Data

E.1 Breakdown of Responses by Question

Table E.1: Average agreement to the statement *This comment is coherent/easy to understand*, by classification present

Annotation	Min.	Lower Quartile	Median	Upper Quartile	Max.	Mean	σ
Information	1.00	2.00	4.00	4.00	5.00	3.475	1.131
Transition	1.00	3.00	4.00	5.00	5.00	3.750	1.112
Logical Attack	1.00	2.00	4.00	4.00	5.00	3.308	1.191
Logical Support	2.00	3.50	4.00	4.00	5.00	3.778	0.737
Rhetorical Attack	1.00	3.00	4.00	4.00	5.00	3.527	1.107
Rhetorical Support	1.00	3.00	4.00	4.00	5.00	3.580	1.113
Preference	2.00	2.50	3.00	3.50	4.00	3.000	1.000
Persona	1.00	3.00	4.00	4.00	5.00	3.511	1.122
Group	1.00	3.00	4.00	4.00	5.00	3.527	1.126
Audience	1.00	3.00	4.00	4.00	5.00	3.510	1.109
Implied Relationship	2.00	3.00	4.00	4.00	4.00	3.429	0.728
Implied Belief	1.00	4.00	4.00	4.75	5.00	3.800	1.108
Spam/Advertisement	1.00	1.00	1.00	1.00	1.00	1.000	0.000
Unknown	1.00	1.75	2.50	3.50	5.00	2.750	1.479
None	1.00	1.25	4.00	4.00	5.00	3.100	1.578

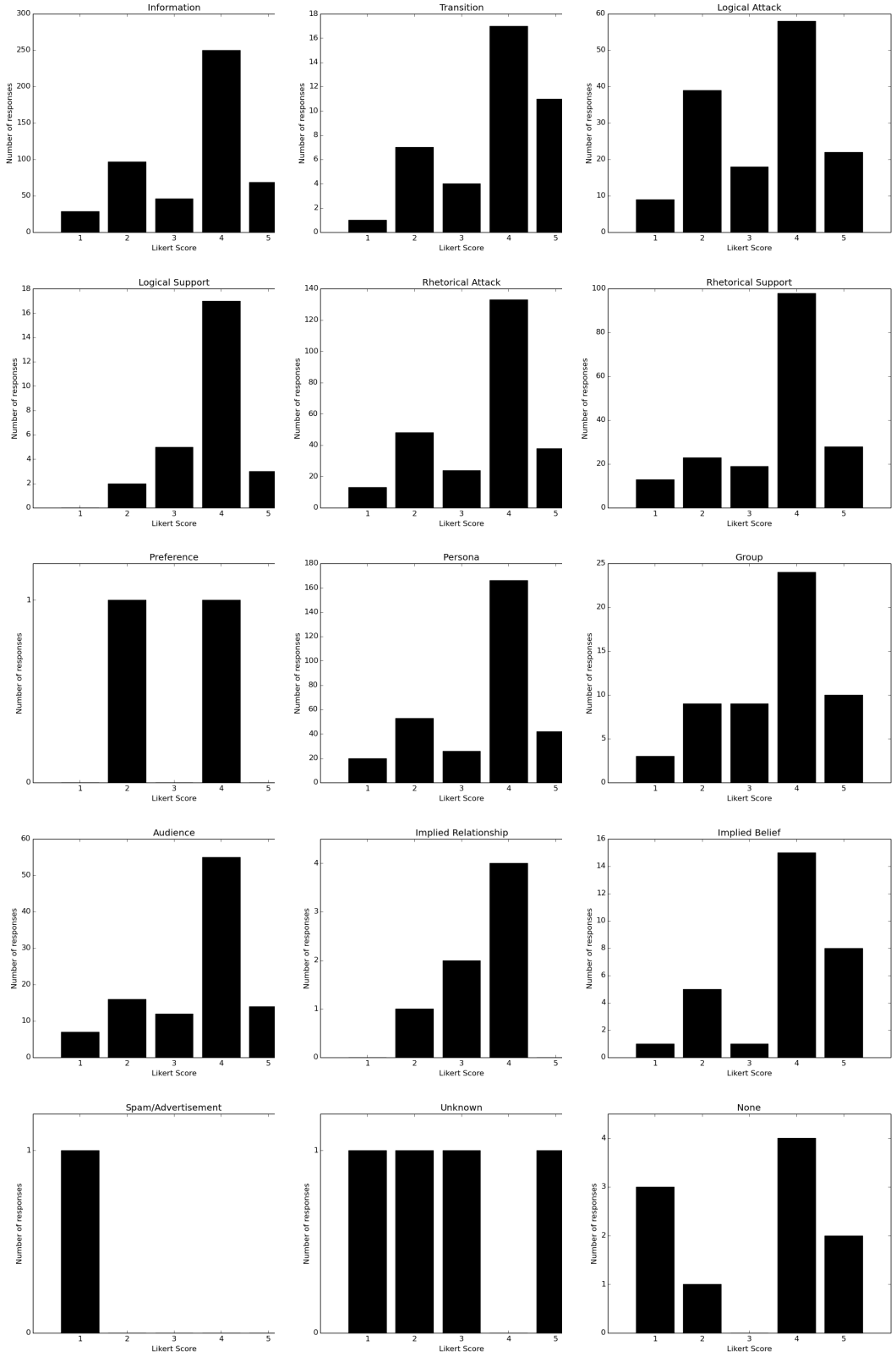


Figure E.1: Distribution for agreement with the statement *This comment is coherent/easy to understand*, by classification present

Table E.2: Average agreement with the statement *This comment contains (or appears to contain) credible information*, by classification present

Annotation	**Min.**	**Lower Quartile**	**Median**	**Upper Quartile**	**Max.**	**Mean**	σ
Information	1.00	2.00	3.00	4.00	5.00	2.756	1.069
Transition	1.00	2.00	3.00	3.00	5.00	2.600	1.091
Logical Attack	1.00	2.00	3.00	4.00	5.00	2.966	1.036
Logical Support	1.00	2.00	3.00	3.50	5.00	2.815	1.020
Rhetorical Attack	1.00	2.00	2.00	3.00	5.00	2.418	1.043
Rhetorical Support	1.00	2.00	3.00	3.00	5.00	2.547	1.084
Preference	2.00	2.00	2.00	2.00	2.00	2.000	0.000
Persona	1.00	2.00	2.00	3.00	5.00	2.427	1.032
Group	1.00	2.00	3.00	3.50	5.00	2.873	1.113
Audience	1.00	2.00	2.00	3.00	5.00	2.356	1.028
Implied Relationship	1.00	2.00	2.00	2.50	3.00	2.143	0.639
Implied Belief	1.00	1.25	2.00	3.00	5.00	2.333	1.135
Spam/Advertisement	1.00	1.00	1.00	1.00	1.00	1.000	0.000
Unknown	1.00	1.00	2.00	3.00	3.00	2.000	1.000
None	1.00	1.00	1.00	2.50	4.00	1.700	1.100

Table E.3: Average agreement with the statement *This comment makes (or attempts to make) a persuasive argument*, by classification present

Annotation	**Min.**	**Lower Quartile**	**Median**	**Upper Quartile**	**Max.**	**Mean**	σ
Information	1.00	2.00	4.00	4.00	5.00	3.196	1.137
Transition	1.00	2.00	3.00	4.00	5.00	2.825	1.138
Logical Attack	1.00	2.00	4.00	4.00	5.00	3.329	1.211
Logical Support	1.00	2.00	3.00	4.00	4.00	3.000	1.089
Rhetorical Attack	1.00	2.00	3.00	4.00	5.00	2.965	1.160
Rhetorical Support	1.00	2.00	3.00	4.00	5.00	2.685	1.192
Preference	3.00	3.25	3.50	3.75	4.00	3.500	0.500
Persona	1.00	2.00	3.00	4.00	5.00	2.886	1.177
Group	1.00	2.00	4.00	4.00	5.00	3.273	1.242
Audience	1.00	2.00	2.50	3.00	5.00	2.596	1.043
Implied Relationship	1.00	2.50	3.00	4.00	5.00	3.143	1.245
Implied Belief	1.00	2.00	3.00	4.00	5.00	3.000	1.155
Spam/Advertisement	1.00	1.00	1.00	1.00	1.00	1.000	0.000
Unknown	1.00	1.00	2.00	3.00	3.00	2.000	1.000
None	1.00	1.00	1.00	2.50	3.00	1.600	0.917

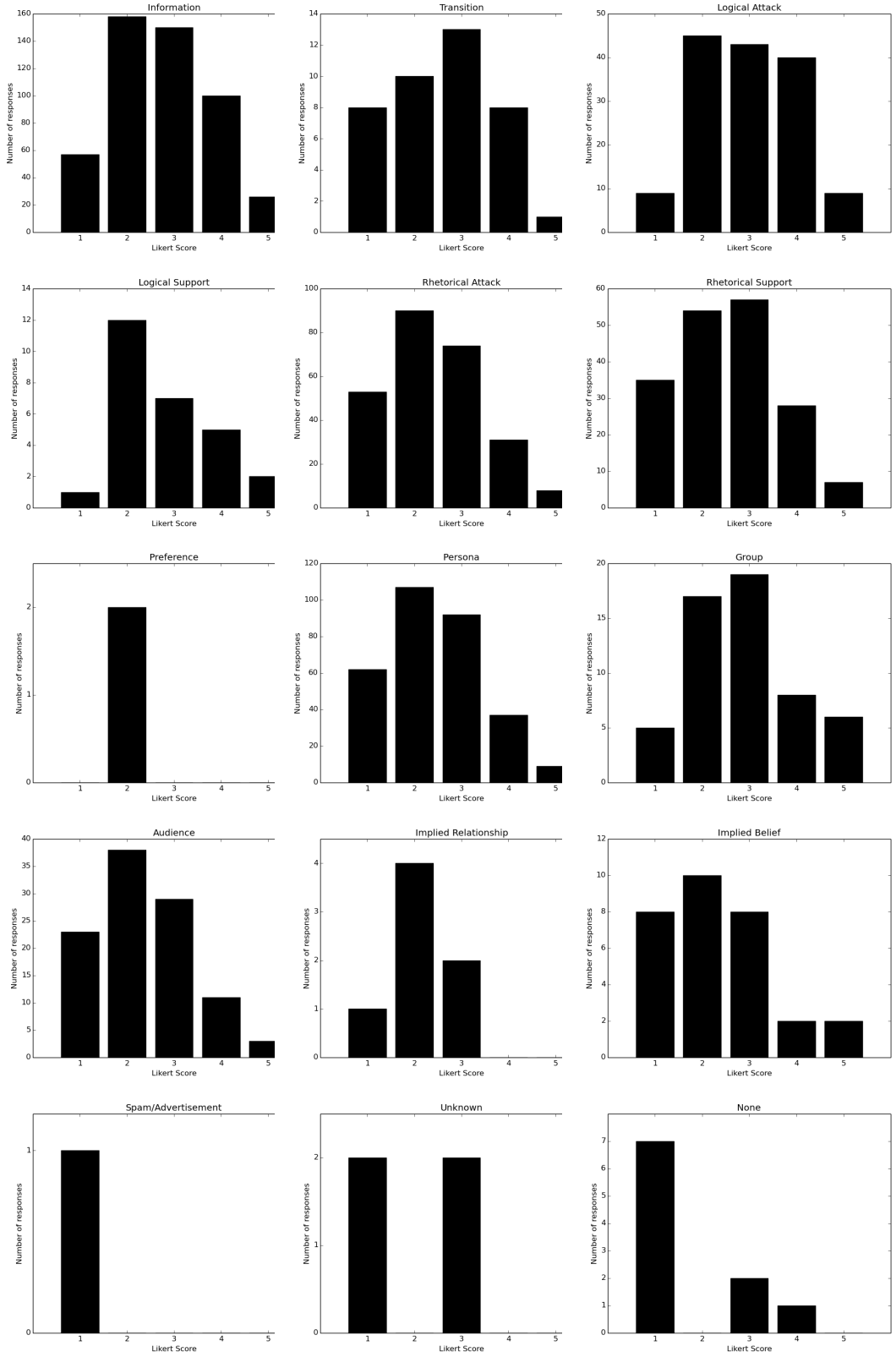


Figure E.2: Distribution for agreement with the statement *This comment contains (or appears to contain) credible information*, by classification present

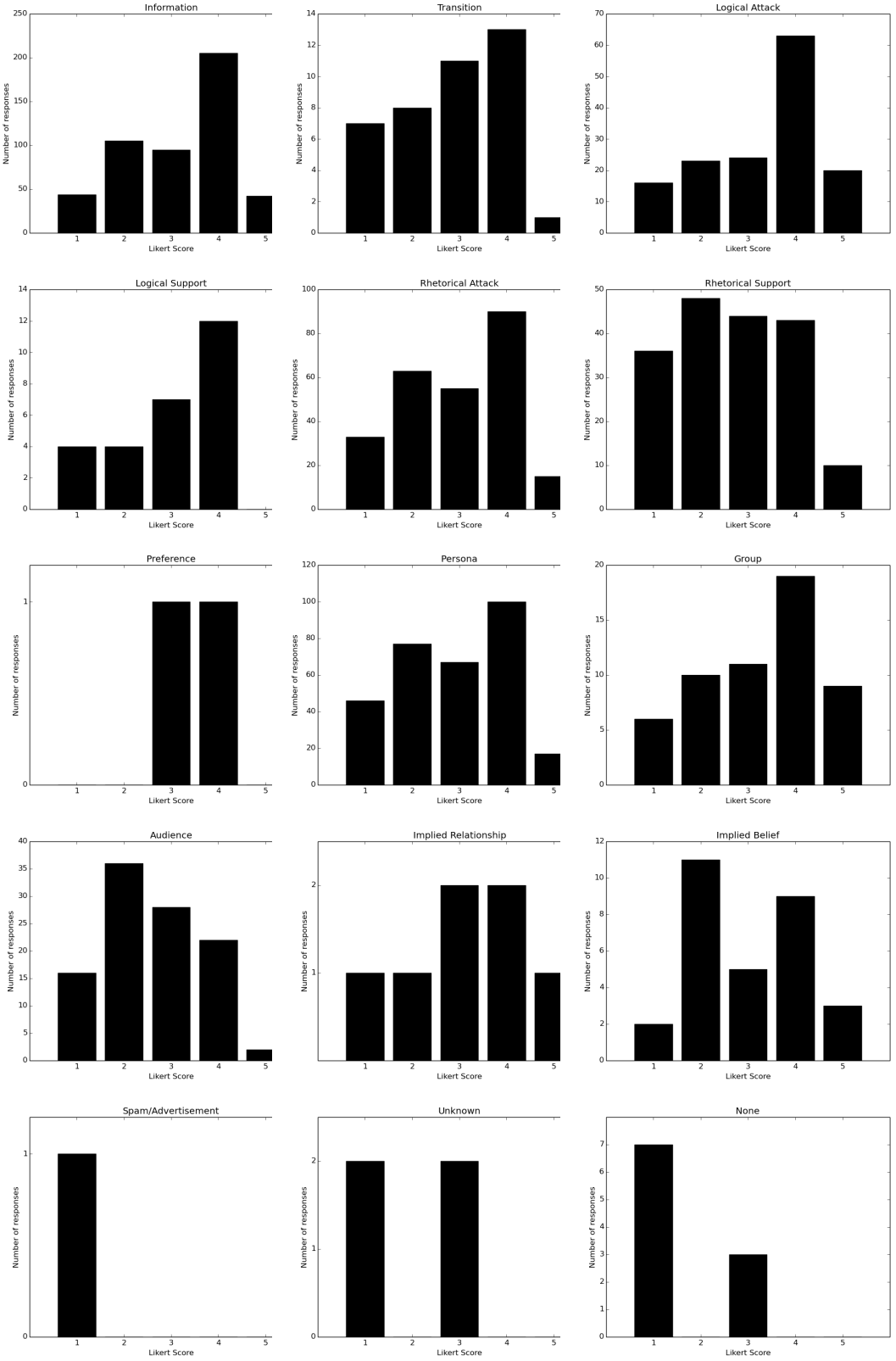


Figure E.3: Distribution for agreement with the statement *This comment is (or attempts to be) entertaining*, by classification present

Table E.4: Average agreement with the statement *This comment is (or attempts to be) entertaining*, by classification present

Annotation	Min.	Lower Quartile	Median	Upper Quartile	Max.	Mean	σ
Information	1.00	2.00	3.00	4.00	5.00	2.678	1.178
Transition	1.00	1.75	3.00	4.00	5.00	2.650	1.216
Logical Attack	1.00	1.00	2.00	3.00	5.00	2.158	1.090
Logical Support	1.00	2.00	3.00	3.00	5.00	2.704	0.974
Rhetorical Attack	1.00	2.00	3.00	4.00	5.00	3.094	1.221
Rhetorical Support	1.00	3.00	4.00	4.00	5.00	3.420	1.142
Preference	3.00	3.25	3.50	3.75	4.00	3.500	0.500
Persona	1.00	2.00	4.00	4.00	5.00	3.280	1.181
Group	1.00	2.00	3.00	4.00	5.00	2.873	1.207
Audience	1.00	3.00	4.00	4.00	5.00	3.740	0.971
Implied Relationship	2.00	2.00	3.00	3.50	4.00	2.857	0.833
Implied Belief	1.00	2.00	3.00	4.00	5.00	2.967	1.303
Spam/Advertisement	1.00	1.00	1.00	1.00	1.00	1.000	0.000
Unknown	1.00	1.75	2.50	3.00	3.00	2.250	0.829
None	1.00	1.50	3.00	4.00	5.00	3.000	1.483

Table E.5: Average agreement with the statement *This comment is (or attempts to be) offensive*, by classification present

Annotation	Min.	Lower Quartile	Median	Upper Quartile	Max.	Mean	σ
Information	1.00	2.00	2.00	3.00	5.00	2.556	1.152
Transition	1.00	1.00	3.00	3.25	5.00	2.575	1.181
Logical Attack	1.00	1.00	2.00	3.00	5.00	2.301	1.088
Logical Support	1.00	2.00	3.00	4.00	5.00	2.963	1.261
Rhetorical Attack	1.00	2.00	3.00	4.00	5.00	3.164	1.141
Rhetorical Support	1.00	1.00	2.00	3.00	5.00	2.309	1.104
Preference	3.00	3.00	3.00	3.00	3.00	3.000	0.000
Persona	1.00	2.00	3.00	4.00	5.00	2.853	1.206
Group	1.00	2.00	3.00	4.00	5.00	2.873	1.192
Audience	1.00	2.00	3.00	4.00	5.00	2.817	1.116
Implied Relationship	1.00	3.00	4.00	4.00	5.00	3.429	1.178
Implied Belief	1.00	2.00	3.50	4.00	5.00	3.333	1.135
Spam/Advertisement	1.00	1.00	1.00	1.00	1.00	1.000	0.000
Unknown	1.00	1.00	1.00	1.50	3.00	1.500	0.866
None	1.00	1.00	1.00	2.75	3.00	1.700	0.900

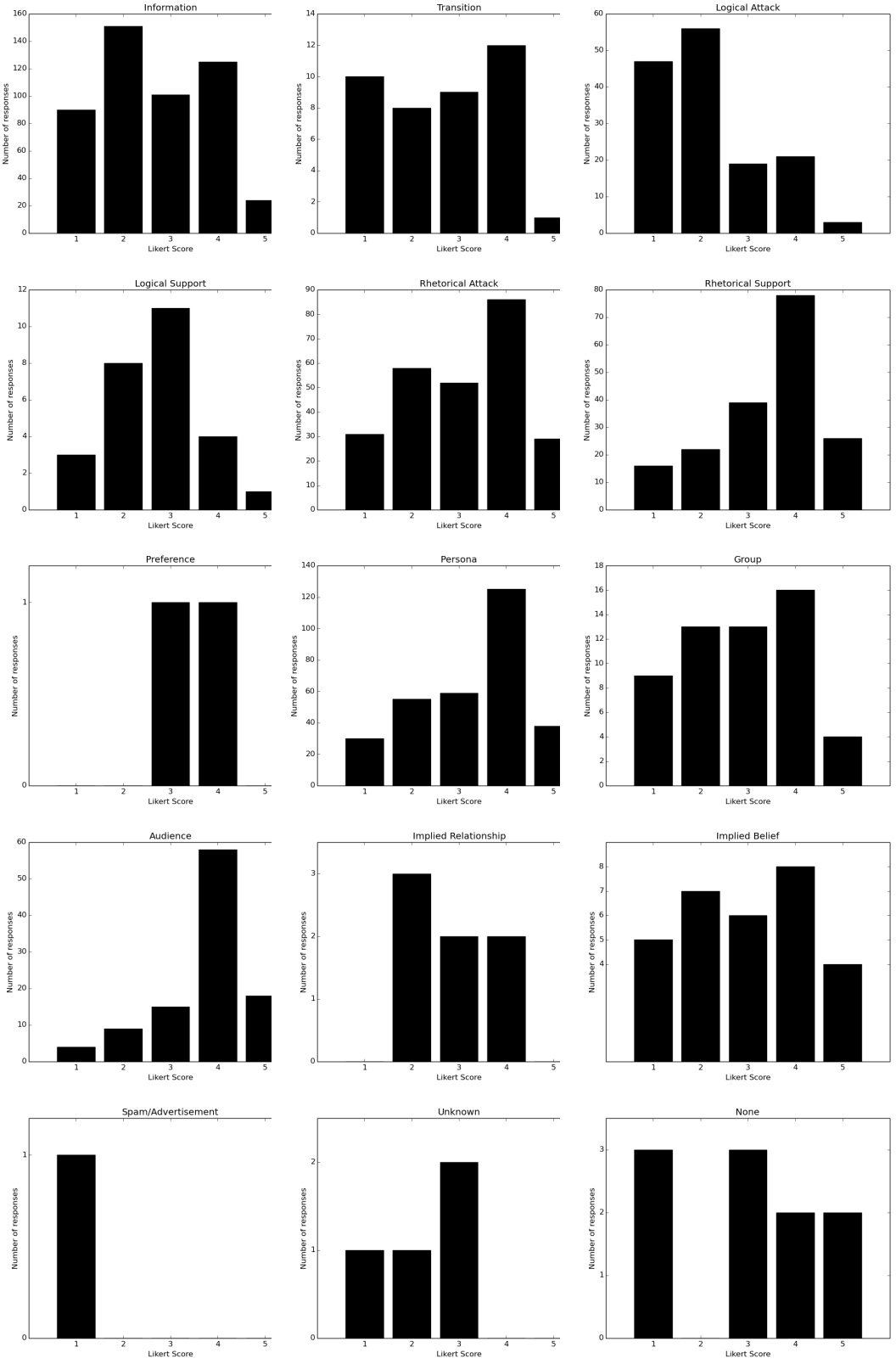


Figure E.4: Distribution for agreement with the statement *This comment is (or attempts to be) entertaining*, by classification present

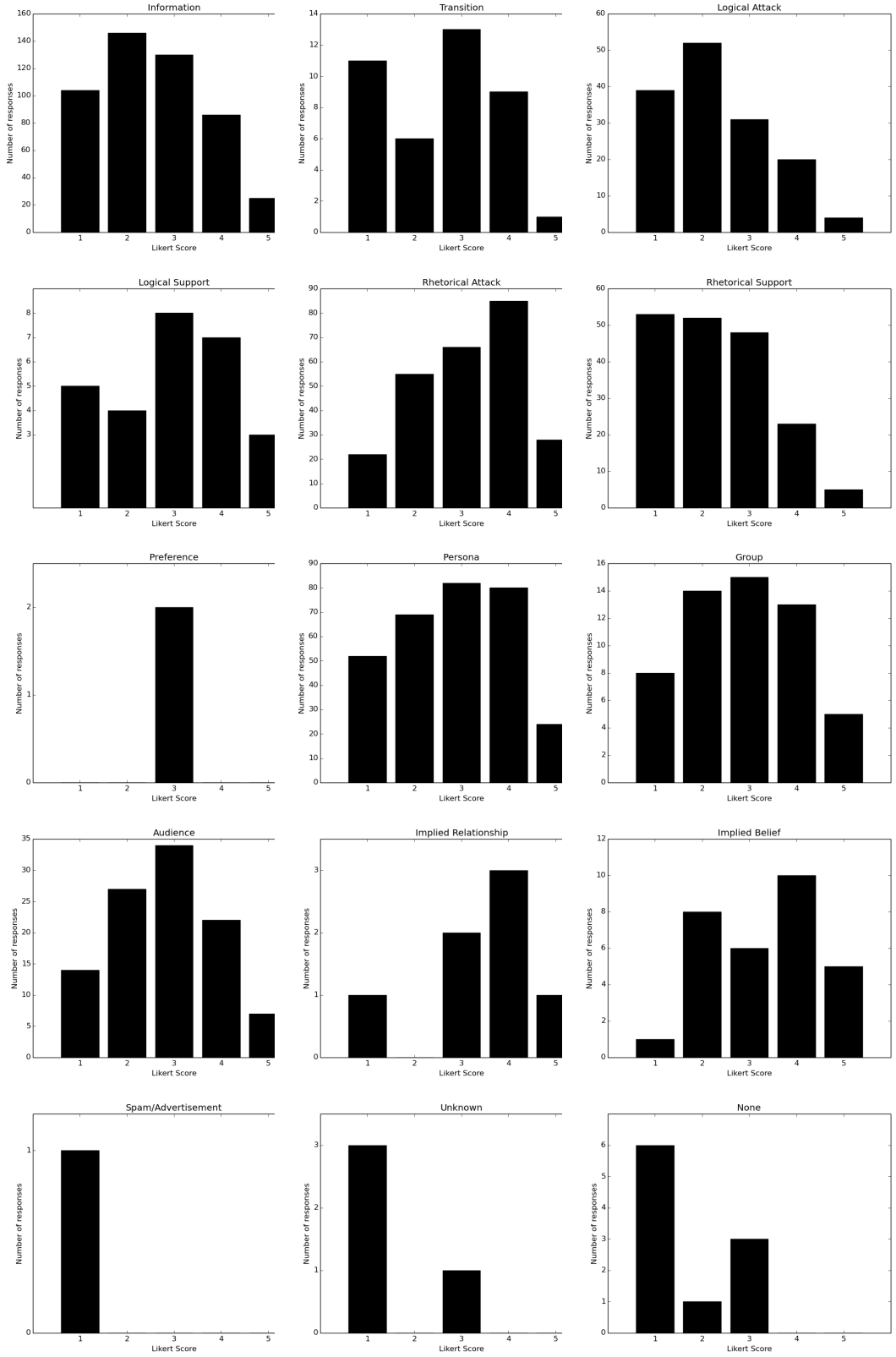


Figure E.5: Distribution for agreement with the statement *This comment is (or attempts to be) offensive*, by classification present

Table E.6: Average response to the question *Would you be more or less likely to reply to this comment than average?*, by classification present

Annotation	**Min.**	**Lower Quartile**	**Median**	**Upper Quartile**	**Max.**	**Mean**	σ
Information	1.00	1.00	2.00	3.00	5.00	2.322	1.088
Transition	1.00	2.00	3.00	3.00	5.00	2.525	1.072
Logical Attack	1.00	1.00	2.00	3.00	5.00	2.390	1.161
Logical Support	1.00	1.50	2.00	3.00	4.00	2.185	0.862
Rhetorical Attack	1.00	1.00	2.00	3.00	5.00	2.199	1.062
Rhetorical Support	1.00	1.00	2.00	3.00	5.00	2.271	1.061
Preference	2.00	2.00	2.00	2.00	2.00	2.000	0.000
Persona	1.00	1.00	2.00	3.00	5.00	2.238	1.052
Group	1.00	1.00	2.00	3.00	4.00	2.200	1.051
Audience	1.00	1.75	2.00	3.00	5.00	2.385	1.059
Implied Relationship	1.00	1.00	1.00	2.00	3.00	1.571	0.728
Implied Belief	1.00	1.00	2.00	3.00	4.00	2.067	0.892
Spam/Advertisement	1.00	1.00	1.00	1.00	1.00	1.000	0.000
Unknown	1.00	1.00	1.00	1.50	3.00	1.500	0.866
None	1.00	1.00	1.00	2.50	4.00	1.700	1.100

Table E.7: Average response to the question *Would you be more or less likely to share this comment (to friends/followers/etc.) than average?*, by classification present

Annotation	**Min.**	**Lower Quartile**	**Median**	**Upper Quartile**	**Max.**	**Mean**	σ
Information	1.00	1.00	2.00	3.00	5.00	2.045	0.950
Transition	1.00	1.00	2.00	3.00	4.00	2.025	0.961
Logical Attack	1.00	1.00	2.00	3.00	5.00	2.048	1.016
Logical Support	1.00	1.00	2.00	3.00	4.00	2.000	0.903
Rhetorical Attack	1.00	1.00	2.00	2.00	4.00	1.863	0.906
Rhetorical Support	1.00	1.00	2.00	3.00	4.00	2.122	1.033
Preference	2.00	2.00	2.00	2.00	2.00	2.000	0.000
Persona	1.00	1.00	2.00	3.00	4.00	2.010	0.970
Group	1.00	1.00	2.00	3.00	4.00	1.927	0.912
Audience	1.00	1.00	2.00	3.00	4.00	2.269	1.058
Implied Relationship	1.00	1.00	1.00	2.00	3.00	1.571	0.728
Implied Belief	1.00	1.00	2.00	2.00	3.00	1.800	0.792
Spam/Advertisement	1.00	1.00	1.00	1.00	1.00	1.000	0.000
Unknown	1.00	1.00	1.00	1.50	3.00	1.500	0.866
None	1.00	1.00	1.00	1.00	3.00	1.200	0.600

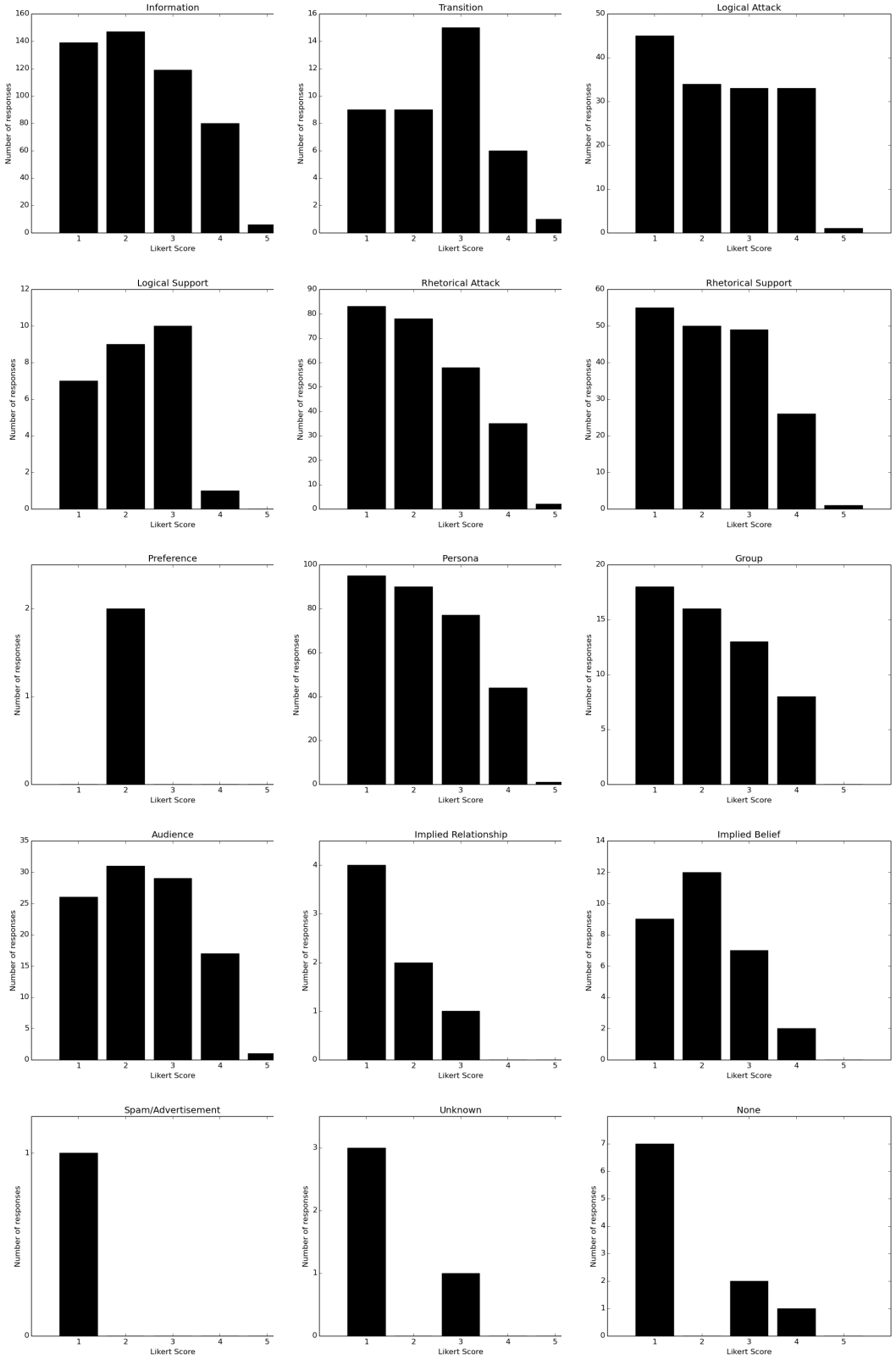


Figure E.6: Distribution for response to the question *Would you be more or less likely to reply to this comment than average?*, by classification present

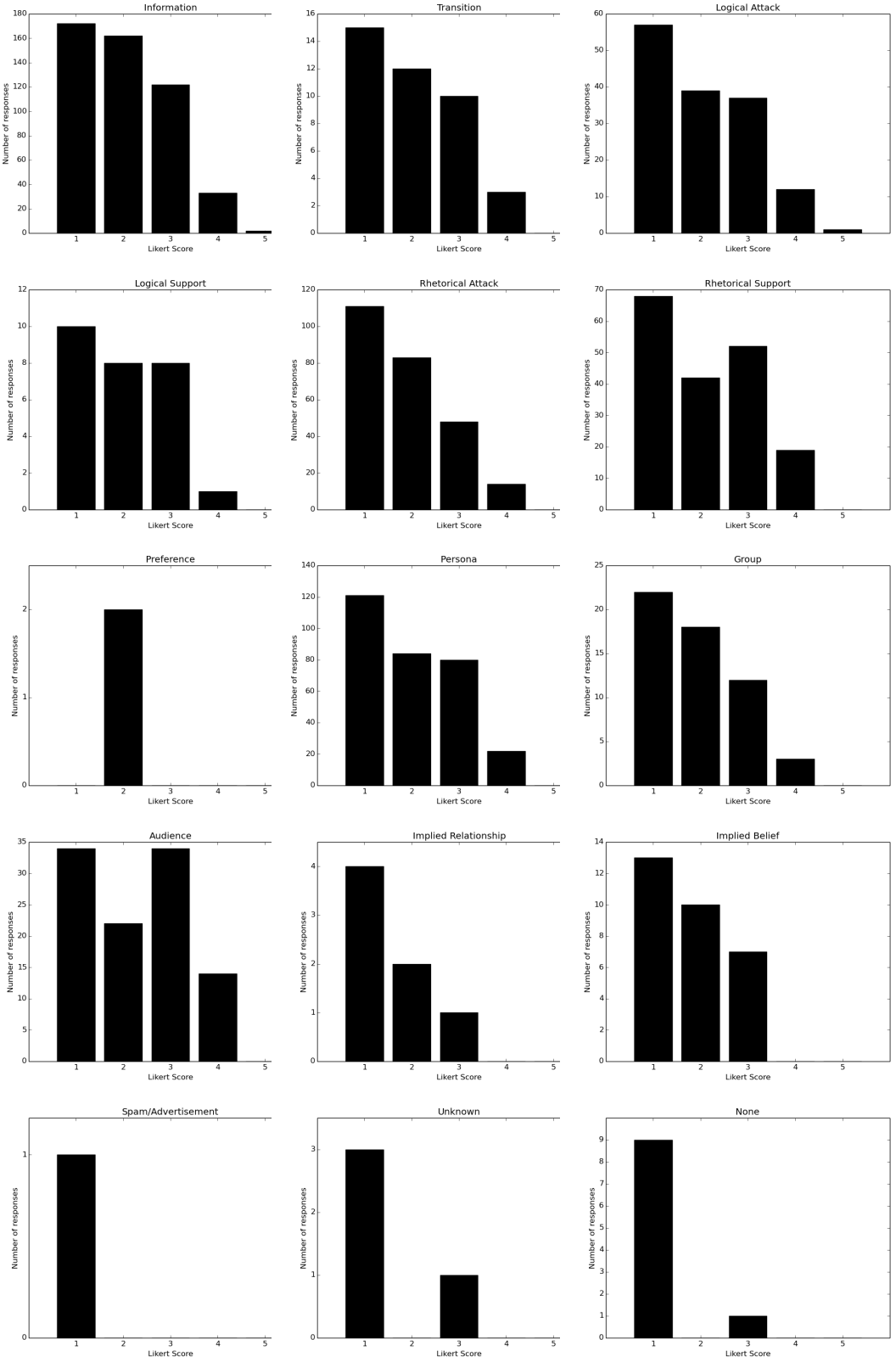


Figure E.7: Distribution for response to the question *Would you be more or less likely to share this comment (to friends/followers/etc.) than average?*, by classification present

Table E.8: Average response to the question *Would you be more or less likely to up-/down-vote this comment than average?*, by classification present

Annotation	**Min.**	**Lower Quartile**	**Median**	**Upper Quartile**	**Max.**	**Mean**	σ
Information	1.00	1.00	2.00	3.00	5.00	2.497	1.162
Transition	1.00	1.00	3.00	3.00	4.00	2.450	1.094
Logical Attack	1.00	1.00	3.00	4.00	5.00	2.548	1.188
Logical Support	1.00	2.00	2.00	3.00	4.00	2.333	0.903
Rhetorical Attack	1.00	1.00	2.00	3.00	5.00	2.426	1.190
Rhetorical Support	1.00	2.00	3.00	3.00	5.00	2.530	1.149
Preference	2.00	2.00	2.00	2.00	2.00	2.000	0.000
Persona	1.00	1.00	2.00	3.00	5.00	2.463	1.167
Group	1.00	2.00	3.00	3.50	5.00	2.618	1.168
Audience	1.00	2.00	3.00	4.00	5.00	2.625	1.145
Implied Relationship	1.00	1.00	2.00	3.50	5.00	2.429	1.498
Implied Belief	1.00	2.00	3.00	4.00	5.00	2.667	1.164
Spam/Advertisement	4.00	4.00	4.00	4.00	4.00	4.000	0.000
Unknown	1.00	1.00	2.00	3.25	4.00	2.250	1.299
None	1.00	1.00	1.00	2.50	4.00	1.700	1.100

Table E.9: Average response to the question *Would you be more or less likely to report this comment than average?*, by classification present

Annotation	**Min.**	**Lower Quartile**	**Median**	**Upper Quartile**	**Max.**	**Mean**	σ
Information	1.00	1.00	2.00	3.00	5.00	1.986	0.932
Transition	1.00	1.00	2.00	3.00	4.00	1.900	0.970
Logical Attack	1.00	1.00	2.00	3.00	5.00	1.973	0.958
Logical Support	1.00	1.00	2.00	3.00	4.00	2.074	0.979
Rhetorical Attack	1.00	1.00	2.00	3.00	5.00	2.094	0.996
Rhetorical Support	1.00	1.00	2.00	3.00	5.00	1.972	0.960
Preference	2.00	2.25	2.50	2.75	3.00	2.500	0.500
Persona	1.00	1.00	2.00	3.00	5.00	2.042	0.999
Group	1.00	1.00	2.00	3.00	5.00	2.182	1.011
Audience	1.00	1.00	2.00	3.00	4.00	2.154	0.948
Implied Relationship	1.00	1.50	3.00	3.50	5.00	2.714	1.385
Implied Belief	1.00	2.00	2.00	3.00	5.00	2.333	1.043
Spam/Advertisement	5.00	5.00	5.00	5.00	5.00	5.000	0.000
Unknown	1.00	1.00	1.00	1.50	3.00	1.500	0.866
None	1.00	1.00	1.00	1.75	4.00	1.600	1.020

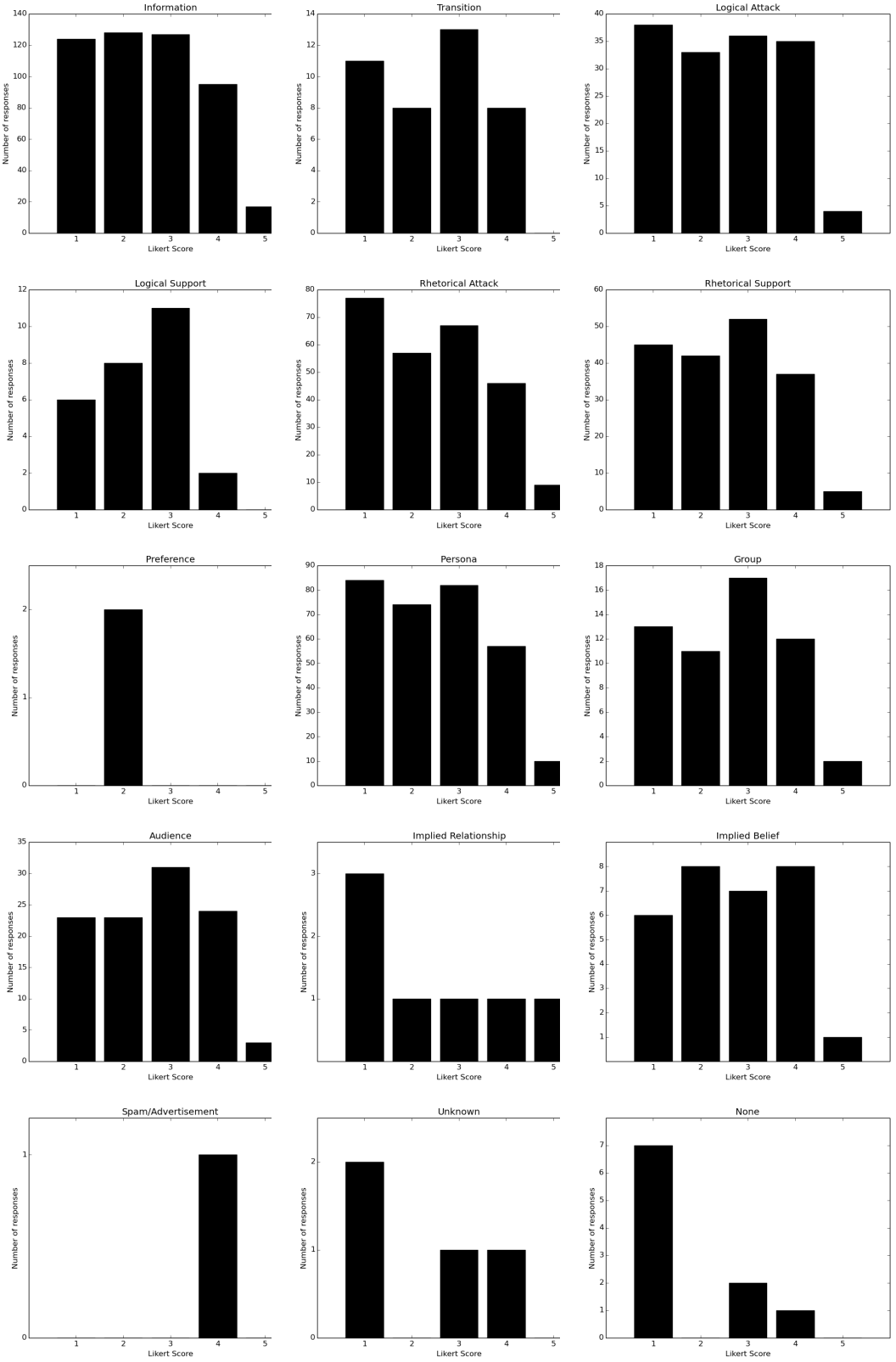


Figure E.8: Distribution for response to the question *Would you be more or less likely to up-/down-vote this comment than average?*, by classification present

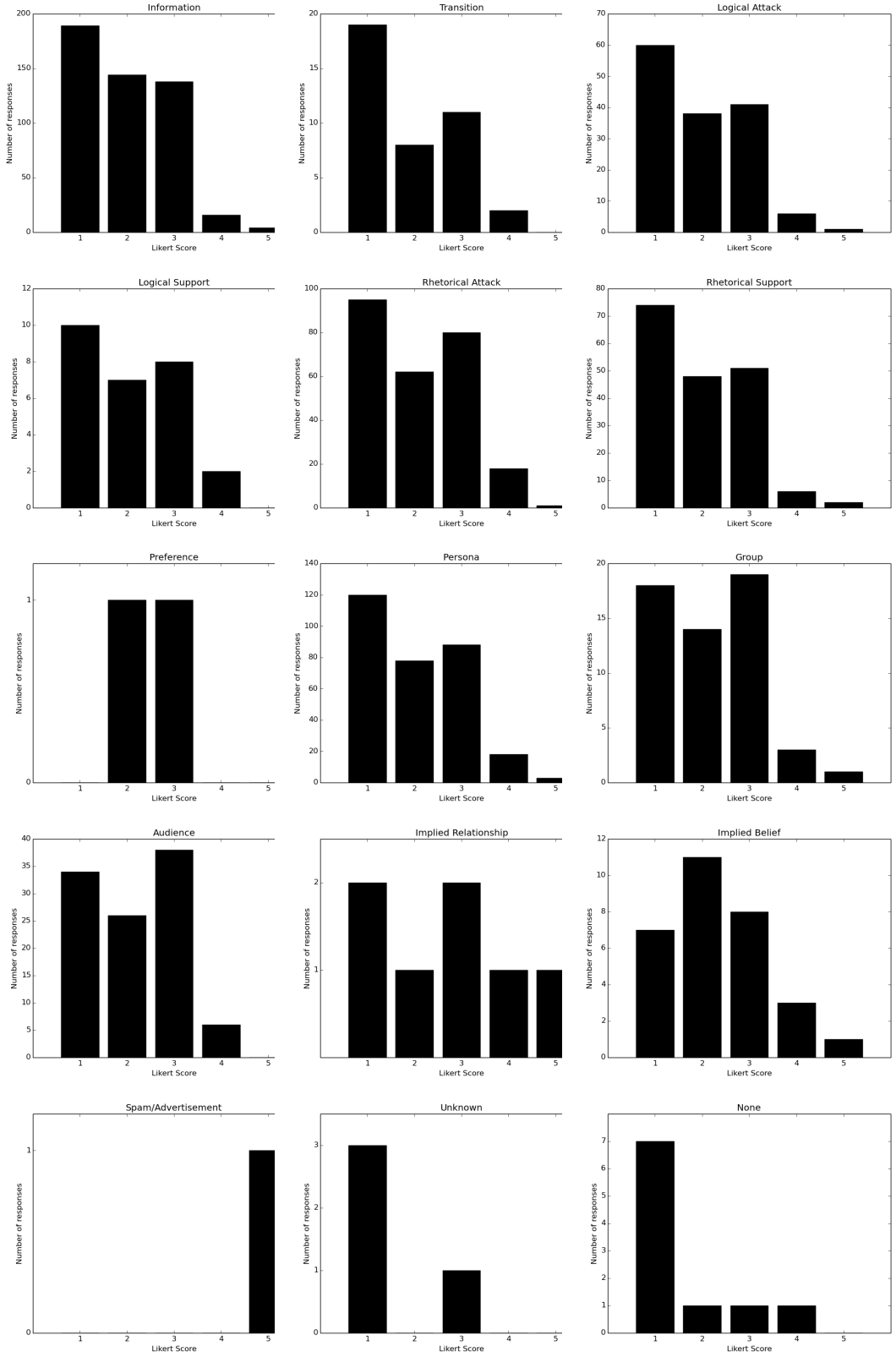


Figure E.9: Distribution for response to the question *Would you be more or less likely to report this comment than average?*, by classification present