What's Wife Swap got to do with it? Talking politics in the net-based public sphere

Graham, T.S.

Citation for published version (APA):
Chapter 4

Research design & methodology

4.1 Introduction

In order to address the aims and questions of this study, a comparative study design with normative, descriptive, and explorative characteristics was utilized. In other words, at one level, this study is a normative analysis whereby the theoretical notions of the public sphere are operationalized into criteria for assessing the deliberativeness of online political talk. While at another level, it is both descriptive and explorative in which an interactional analysis on the pragmatic and functional components of political talk is carried out as a means of providing a more accurate picture of how the political emerges in online discussions, how people actually talk politics in those discussions, and finally, how non-traditional deliberative communicative forms such as humor, emotional comments, and acknowledgements interact and influence the more traditional elements of deliberation. These analyses are conducted within the framework of a comparative study design of three online discussion forums: a politically oriented forum, a non-politically oriented forum, and a mixed forum. As a means of investigating and analyzing political talk within these forums, a content analysis with both qualitative and quantitative features was employed as the primary instrument for examination.

In this chapter then, the research design and methodological approach are laid out in detail. In section 4.2, the design is discussed at length. In particular, the type of design, along with the rationale for selecting it, is given. In section 4.3, the data collection procedures are explained. Specifically, the sampling, archiving, organizing, and managing of the data, along with ethical considerations are provided. In section 4.4, the methodological approach is mapped out, specifically, a guide on how the data were collected, analyzed, and assessed. Finally, in section 4.5, a discussion on the limitations of this study is provided.

4.2 The research design

As discussed earlier, there are two basic problems with net-based public sphere research to date. First, these studies have focused solely on politically oriented discus-

\[^{37}\text{An earlier version of the methodological approach was published in (Graham, 2008) Javnost—The Public 15(2), 17-36.}\]
sion forums. However, political talk is not exclusively reserved for these types of forums. From fan-based forums to numerous other forum genres, discussions on an array of political topics emerge throughout the online communicative landscape, which also contribute to the web of informal conversations that constitutes the public sphere. Second, when analyzing political discussion forums, these studies have tended to privilege a formal notion of deliberation. They have focused mostly on examining the use of rationality via argumentation and have ignored emotions and other communicative forms, typical ingredients of everyday informal political talk. Thus far, net-based public sphere research has only provided us with a partial picture.

In this section, the research design aimed at addressing these problems is laid out. In particular, the normative, interactional, and comparative approaches are discussed. Finally, this section ends with a brief account of the three forums selected for this study along with the rationale for their selection.

4.2.1 Normative analysis of political talk

Moving beyond politically oriented forums and beyond rationality does not mean an abandonment of a normative approach to deliberation. On the contrary, at one level, this is still a normative study. One of the central aims is to assess online political talk in light of the normative conditions of deliberation as outlined in Chapter 2. The first central research question of this study is: To what extent do the communicative practices of online political discussions satisfy the normative conditions of the process of deliberation of the public sphere? In order to address this question, a normative evaluative approach was integrated with empirical investigations. Dryzek (1995) identifies such an analysis as ‘pure-critical approach’, evaluating and criticizing real-world practices to the extent they fall short of the ideal. Dahl (1967) maintains that such an approach requires formulating explicit norms that are operationalized into measurable concepts, which are then applied to an empirical analysis. In other words, such an approach requires specifying, operationalizing, and applying the normative conditions.

In order to carry out the normative analysis, these three steps were employed. In Chapter 2, the normative conditions of deliberation were specified. Here a set of democratic ideals on the process of deliberation from deliberative democrats and public sphere theorists was constructed. Later in this chapter, these conditions are operationalized into empirical indicators thereby creating the necessary tools for assessing political talk. Finally, in the results chapters, these indicators are applied to political talk from three different types of discussion forums. Here the normative criteria are applied universally to all three forums.

4.2.2 Interactional analysis of political talk

As stated above, this study looks to move beyond a normative notion of deliberation. The aim is to provide a more authentic account of how people actually talk politics
Research design & methodology

online, and to see how other communicative forms such as expressive speech acts interact and influence the variables of deliberation. The second central research question being addressed here is both descriptive and explorative: what role, if any, do expressives play within online political discussions and in relation to the normative conditions of deliberation? In order to address this question and to provide a more comprehensive account of online discussions, an interactional analysis on the pragmatic and functional components of political talk was employed.

Such an analysis was chosen for several reasons. Given that political talk is a social process, a conversation, an interactional approach was chosen because it treats texts, speech acts, as part of a social process. As Nofsinger (1991) maintains, speech acts are not separate individual actions, but rather they are integrated components in the ongoing flow of a conversation. Moreover, a pragmatic approach places importance on participants’ practical operation of political talk in actual communicative situations. Consequently, such an approach not only views political talk functionally, but more importantly, it treats it as a form of strategic action by participants in context. As Nofsinger (1991, p. 7) argues, “Participants are not merely saying something to each other when they talk. They are doing something at the same time: directing communicative or social actions at one another”. Finally, such an analysis is also an effective means of operationalizing and carrying out (some of) the normative components of this study.

4.2.3 Comparative analysis of political talk

In order to address the shortcomings discussed above, the normative and interactional approaches were conducted via a comparative study design. A comparative approach was chosen because, simply put, it provides an effective and productive means for carrying out one of the central aims of this study, which is to move beyond the political. In particular, it allows for a comparison of political talk between and across political, nonpolitical, and mixed forums so that a more fruitful and insightful examination and investigation may be conducted. Moreover, by including a political forum, the analysis is constantly presented with a reflection of the so-called ‘political’, consequently, strengthening it and presenting it with a strategy for explanation. Finally, through comparison, the differences between the three forums that emerge may provide additional insight into the individual forums.

It should be noted here that the comparative analysis is presented in a cumulative fashion in the following three results chapters. In Chapter 5, the findings from the Guardian are discussed in comparison to past studies. In Chapter 6, the findings from the Big Brother forum are discussed in comparison to the Guardian. Finally, in Chapter 7, the Wife Swap findings are discussed in comparison to both the Guardian and Big Brother. This format was chosen because it proved to be the most effective and efficient means of presenting the comparative analysis; it provided enough space for the individual cases while avoiding unnecessary repetition.
The forums

There were four criteria applied to selecting the forums for this study. The first criterion focused on finding forums that would best illustrate a dichotomy between the political and the nonpolitical in a traditional sense, between the so-called high and popular cultures; places where one might expect to find ‘serious’ political talk and ‘not so serious’ talk. Second and third criteria consisted of selecting forums from the same national context in order to provide a more fruitful analysis between forums and selecting forums based on popularity in terms of both name recognition and forum traffic/participation. Finally, forums were selected based on language; English speaking forums, the author’s mother tongue, were selected. Based on these criteria, three discussion forums were selected: the Guardian’s Politics Talkboard, Channel 4’s Wife Swap Forum, and the Big Brother Fan’s Celebrity Big Brother Forum.

The Guardian. The Guardian is a British newspaper, which is owned by the Guardian Media Group. Its talkboards are hosted by the website guardian.co.uk, which represents its online presence. It is one of the leading online newspapers in the UK and contains nearly all the content from its offline counterparts (the Guardian and the Observer) along with its own original material. The Guardian represents a ‘quality’ newspaper, and one would expect its talkboards to host ‘serious’ political talk. An exploratory study (Graham & Witschge, 2003) revealed that indeed the discussions that took place within the Guardian’s Politics Talkboard were deliberative. Moreover, it hosts a multitude of participants and discussions on a diverse range of national, European, and international political topics.

Big Brother Fan. The Celebrity Big Brother discussion forum is hosted by bbfans.com, which is a website ran by and dedicated to fans of the reality TV series Big Brother UK, which is broadcasted by Channel 4. The website offers a variety of forums on Big Brother, Big Brother spinoffs, reality TV, and on other entertainment oriented topics and media. Moreover, the forums are lively communicative spaces; they maintain thousands of participants, which have contributed hundreds of thousands of postings. The specific forum selected for this study was Celebrity Big Brother 2006.

Celebrity Big Brother, which first aired in 2001, is now a full spinoff of Big Brother UK. The series features a number of celebrities living in the Big Brother House, who try to avoid eviction by the public with the aim of winning a cash prize to be donated to the their nominated charity at the end of the series. What makes Celebrity Big Brother 2006 interesting is that one of the housemates in that series was the British MP, George Galloway. Galloway maintained publicly that his

---

38 Talkboard is another word for discussion forum.
39 Available at: http://politicstalk.guardian.co.uk/WebX?:14@@.ee80025
40 George Galloway is a former Labour MP who was expelled from the party due to his outspoken comments on the Iraq War. He currently is a member of the Respect party and represents the Bethnal Green and Bow constituency.
appearance on the show would be good for British politics. In particular, he stated that one of the primary reasons for entering the Big Brother House was to reach out to young people.\footnote{The interview is available at: http://news.bbc.co.uk/1/hi/uk_politics/4587448.stm} Celebrity Big Brother offered a unique communicative space; a nonpolitical oriented discussion forum influenced by a political personality. Consequently, in this study, it represents the mixed forum.

Wife Swap. The Wife Swap discussion forum is hosted by the British public-service television broadcaster Channel 4 and is tied to their TV series Wife Swap.\footnote{As they state, “The Channel’s primary purpose is the fulfilment of its public service remit, which was most recently defined in the 2003 Communications Act.” Available at: http://www.channel4.com/about4/overview.html} Wife Swap is a reality television program produced by the British production company RDF Media and first aired in 2003. The premise behind the show is that in each episode, two families, usually with different family lifestyles and from different social classes, swap wives (mothers) for two weeks. The show has been a success both in the UK and internationally with various versions of it now appearing throughout the world.

The Channel 4 website hosts a community space dedicated to the programs they broadcast. This space offers a variety of discussion forums where fans can discuss together their favorite, or not so favorite, TV shows. These forums are usually filled with participants and discussions. The Wife Swap forum is located within this space under the entertainment category.\footnote{Available at: http://community.channel4.com/groupee/forums/a/cfrm/f/31060416} According to the forum, it is supposed to provide a communicative space where fans can “chat about Wife Swap”. Consequently, Wife Swap represents a nonpolitically oriented forum tied to a reality TV series. In other words, it represents the other side of the dichotomy, a fan-based popular culture forum genre, a place where one might expect to find ‘not so serious’ political talk.

4.3 Data collection procedures: Sampling, archiving, and organizing

In this section, a detailed account of the data selection process is outlined. In particular, the sampling, the archiving, and the organization and management of the data are presented. Additionally, ethical considerations are discussed.

4.3.1 The initial sampling

The research approaches discussed above focus on examining online political discussions. In particular, they focus on the communicative practices of forum participants, as they are externalized in or can be externalized from the postings, and on the
interaction (or lack thereof) between those postings within a discussion thread. Therefore, the unit of analysis, which will be discussed later in this chapter, was the posting. The context unit of analysis was the thread in which the posting was situated. Consequently, as a means of keeping the social integrity of the discussions, the unit of selection was a discussion thread as opposed to the individual posting.

Given the diversity of the three forums discussed above, particularly the size and orientation of the forums, the sampling criteria varied between all three cases. First, since both the Wife Swap and Big Brother forums were tied to a reality TV series, the selection of the threads was based on the broadcasting premier of that particular series. This period was selected because it was thought to be the most active and most relevant time frame within these forums. Thus, the first criterion was the broadcasting dates of the series.

In Wife Swap’s case, series five was chosen, which originally aired between January and March 2005. Consequently, all those threads, which began on or between 1 January and 31 March 2005 were selected. In all three forums, messages were accompanied by a posting date, thus, a verification of the birth of a thread could be easily carried out by checking the date of the initial posting. The archiving of the Wife Swap forum was performed on November 11, 2005. Note that all postings within these threads at the moment of archiving were included in the sample and not just those posted between January and March. The initial sample for Wife Swap contained 79 threads consisting of 892 postings.

In Big Brother’s case, series four of Celebrity Big Brother was chosen due to the presence of George Galloway. The show originally aired during the month of January 2006. Consequently, all those discussion threads, which began on or between 1 and 31 of January 2006 were selected. The archiving of the Celebrity Big Brother forum was performed on March 13, 2006. Note again that all postings within these threads at the moment of archiving were included. The initial sample for Big Brother contained 345 threads consisting of 6803 postings.

The Guardian presented a different situation than the above forums. It was not tied to a television series, and it was the largest of the three forums in terms of threads and postings. Moreover, the forum was divided into 12 sub-forums based on broad topics of discussion. However, at the time of archiving, only four of these sub-forums had more than 15 threads, consequently, they represented the bulk of the activity within the Guardian. The sub-forum selected was called In Britain. It was selected because of its size and the relevance of the topics discussed. Regarding relevance, it was the only sub-forum left that dealt with domestic politics. The assumption was that the topics within the other two forums would also be domestically oriented. Consequently, this sub-forum was selected for comparative reasons.

However, the size of this sub-forum was still too large to serve as the sample. Thus, after an initial review of the posting rate and distribution over time within this sub-forum, a one-month period was selected as the final criterion. All those discussion threads, which began between 1 and 31 of May 2006 were selected. The archiving of the Guardian forum was performed on July 2, 2006. Note again that all
postings within these threads at the moment of archiving were included. The initial sample for the Guardian contained 37 threads consisting of 1271 postings.

4.3.2 Archiving, organizing, and managing the data

Both Big Brother and the Guardian were archived using the software program HTTrack. The program allowed the discussion forums to be downloaded from their site to a local hard drive. In particular, it archived recursively all directories along with HTML, images, and other files from the corresponding server. Thus, the program creates a mirrored website of the original, which allows the user to browse the selected site from link to link as if viewing it online.

Once the discussion forums were archived, selected discussion threads were transferred to MAXQDA. MAXQDA is a software program, which supports textual and content analyses. In this case, MAXQDA was primarily used as a means of organizing and managing the data. In particular, it was selected because it is an effective and efficient means of coding the data, tracking coding decisions, and retrieving the data. There was one feature, the code relation browser, which on occasions was used to assist in the analysis. This browser identifies any relationships between codes that emerged (intersections between codes). Finally, in addition to MAXQDA, both SPSS and Pajek programs were used to assist in various analyses and presentations of the data.

4.3.3 Ethical considerations

Ethical considerations were well thought-out before the collection of the data. Though online research still represents a fairly new phenomenon, there has been a growing body of literature on online research methodologies, in particular, on ethical guidelines for conducting such research (see e.g. Jankowski & Van Selm, 2007). The main concern facing this study, regarding the collection of data, was whether consent was required to quote and analyze the postings from the above forums. According to Herring’s (2001) guidelines for conducting a computer-mediated discourse analysis, informed consent is always required when researching private computer-mediated communication (CMC), while it is not required when researching public CMC, with two exceptions: when the researcher interferes with the subjects or when the subject’s real identity is used.

In this study, these guidelines were followed. The data collected and used came from three public discussion forums; these discussions were not password protected. In other words, any individual with an internet connection and the right URL address

44 Wife Swap was archived by copying the threads directly to a word document, after which, the political threads were transferred to MAXQDA. HTTrack is available at: http://www.httrack.com/
45 Only the political threads were transferred.
46 Pajek is a social network analysis program, and SPSS is a statistical analysis program.
could access them. Moreover, as will be discussed below, the research conducted here was non-obtrusive. In other words, no interaction or interference took place with the participants. Moreover, no personal information about the participants’ identities was acquired, and special care was taken to remove all forum call signs (nicknames) from the texts and replace them with invented ones.47

4.4 Identifying, describing, and assessing political talk

In this section, the various stages and phases of the methodological approach are outlined and discussed in detail. It consisted of two stages of analysis. During the first stage, the initial sample, as reported above, was analyzed for the presence of political talk. Those threads which contained a political discussion were advanced to stage two. During the second stage, the communicative practices of and between participants within the political threads were examined. It is important to note that during this stage all the postings within the threads were included in the analysis, not just the political exchanges. This decision was made because it simply proved too difficult to disentangle the political postings from the nonpolitical ones without losing elements of those debates, thus jeopardizing the integrity and quality of the analysis. Moreover, excluding these postings would contradict in some ways the aims of this study.

In order to conduct these stages of analyses, a qualitative content analysis (Mayring, 2000) was adopted as the primary instrument for investigation and examination. It was deemed the most appropriate method for several reasons. First, a content analysis was selected because it is non-obtrusive; it provides a means of studying naturally occurring discussions in an online setting. Second, as Wilhelm (1999, p. 163) argues, “It is not necessary to know who the participants are, from what walk of life they come from or with what political parties they are affiliated, to paint a compelling portrait of the deliberativeness of these discussions.” Moreover, messages in conversation represent steps in participant’s communicative and social strategic plans in context, which exhibit recurrent patterns that can be analyzed in detail via an analysis of the text (Nofsinger, 1991). Finally, given the diverse nature of the various variables of deliberation and political talk in general, this type of content analysis was deemed most suitable because it allowed for various levels of operationalization, interpretation, and maneuvering.

In the remainder of this section, this two-staged methodological approach is laid out in detail. During stage one, a set of criteria for identifying a political discussion within a text is given. Moreover, the method for identifying the triggers of political talk is also discussed. During the second stage, the coding scheme for describing, exploring, and assessing political talk is specified, and the coding categories are defined. Next, the coding scheme in relation to the normative conditions of

47 I am referring to the forum participants. The identities of the Big Brother housemates were used because they are celebrities.
deliberation—how the coding categories form indicators (the operationalization)—is discussed. Finally, the procedures used for examining and exploring the use of expressive speech acts are presented.

4.4.1 Stage one: Identifying (the triggers of) political talk

As discussed in the introduction, there is a need for a porous approach to what is political, an approach e.g. that allows for a politics of sexuality, of health, of the body, of childcare. Politics today has become more pervasive, and as such, any concept of what is political must be capable of capturing an increasing number of issues and concerns. This is particularly true when exploring the nonpolitically oriented online discursive landscape. So then, what is political? More specifically, how do we identify within a text a political discussion?

Mansbridge’s (1999, p. 214) definition of political is a solid starting point here. For her, a political discussion emerges when a participant draws attention to something that he or she thinks the public should discuss collectively. Under this account, seemingly private issues can emerge as political so long as there are reasons given as to why this should be a collective concern; naturally, these issues can be contested by others. Moreover, such issues do not have to be connected to institutional politics, nor do they require a response from the state. Additionally, action, which has been commonly tied to the notion of political, need not be the result of talk outside the action of talk itself.

Based on this understanding, two criteria for identifying when a discussion turns political within a text were composed. During the first stage of analysis, all discussion threads from all three forums were subjected to these criteria. All those threads, which contained a posting where (1) a participant makes a connection from a particular experience, interest, issue, or topic in general to society, which (2) stimulates reflection and a response by at least one other participant were advanced to stage two of the analysis. It should be stressed here that the aim of the criteria as a whole was to identify a political discussion. The criteria will now be applied to a discussion thread from the Wife Swap forum as a means of demonstrating them in-use:

**Elizabeth:** I think Wife Swap is a good show to educate Jo Public. It shows different families and different ways of parenting. We learn.

**John:** Educate the public in what exactly? Do we need educating on how other families live? We all have friends’n’family members that live completely different to us... we KNOW everyone’s different. Sorry... but...educating Wife Swap aint.

---

48 A posting can be identified as political if it meets the first criterion, but if it fails to meet the second criterion, it is not part of a political discussion. It would be interesting to examine and compare those instance when a topic has been successfully politicized by fulfilling both criteria with those instance when a topic has failed, achieving only the first.
Mary: I think many people do, yes. If people were more educated about other cultures/sub cultures then maybe there would be a little less prejudice and blind hatred in this world.

First, when individual experiences, issues, or topics in general are discussed, there needs to be a connection made from that instance to society. The word connection here implies that the experience, issue, or topic under discussion should be considered as a collective concern and as such discussed collectively. In this example, participants are discussing whether Wife Swap is a good educational tool for society. Elizabeth’s first statement represents the connection from an experience to society, and her second statement qualifies why. Her posting implies that Wife Swap is a good TV series for the public because people learn about different families and different ways of parenting.

The second criterion operationalized the social aspect of political talk. The process of deliberation is a social process. It requires reciprocity and reflection; participants must listen, reflect, and respond to each other. Thus, once the connection is made, it must stimulate reflection among and a response by other participants. The response should question, contest, affirm, or elaborate on the connection. Both John and Mary’s statement fulfill the second criteria, though, in different directions. John contests Elizabeth’s position by arguing that the public needs no education and Wife Swap is not the place. Mary, on the other hand, not only states an affirmation but also takes a step further by suggesting that if people were educated about different cultures, they would be less likely to be prejudice.

**Triggers of political talk**

*How does political talk emerge in nonpolitical discussion forums?* In order to address this final central research question, an examination and investigation aimed at identifying the triggers of political talk within both Big Brother and Wife Swap was conducted. As discussed above, the initial postings, which began the political discussions, were identified. Consequently, a closer reading of the postings leading up to political talk was made possible and as such carried out. However, prior to the analysis (also during), additional measures were taken to improve it. In particular, both Celebrity Big Brother Highlight episodes and Wife Swap episodes and links to third-party sources within the particular postings were consulted when applicable as a means of providing more context to the discussions in question.

In order to conduct this analysis, Mayring’s (2000) procedures for carrying out the development of inductive coding categories were employed. Given that an initial reading of the political threads had already been conducted, a set of tentative triggers was initially developed. After which, three additional rounds of reading and working through the selected material were carried out. During this time, triggers were modified, combined, removed, and new ones created via feedback loops. Additionally, several patterns were identified in relation to the triggers. For example, certain triggers tended to be an overflow of political discussions that were already occurring
in, for example, the Big Brother House, while other triggers seemed to ignite original political discussions in the forums themselves. That said, after the third round, a set of main triggers were deduced.

### 4.4.2 Stage two: Describing and assessing political talk

The coding scheme presented below was developed as a means of analytically describing and normatively assessing how participants talk politics. It moved beyond a formal notion of deliberation and allowed for a more comprehensive description of political talk, allowing emotions and other communicative forms a place in the analysis. Normatively, it provided the tools for a thorough evaluation and examination of the quality of debate. It consisted of three phases.\(^{49}\)

During the first phase, the coding categories were divided into two groups, which aimed at identifying the message type. The two group headings were `initial` and `response`. The unit of analysis during this phase was the individual message. Once all messages were coded, phase two of the scheme began; messages that provided reasoned claims were advanced. During the second phase, the coding categories were divided into two groups: `evidence type` and `argument style`. Messages were first coded for the type of evidence used, after which, selected messages were coded again for argument style. The unit of analysis during this phase was the argument. During the final phase, the coding categories were divided into four groups: `communicative empathy`, `discursive equality`, `discursive freedom`, and `sincerity`. All messages here were coded for various variables of deliberation. The unit of analysis again was the individual message.\(^{50}\) For all three phases, the context unit of analysis was the discussion thread; the relationships between the messages within a single thread were analyzed. The individual coding categories are defined and discussed in detail below.\(^{51}\)

**The coding categories**

**Phase one.** The goal of the first phase of analysis was to identify the message type. Here, messages were coded as one or more of two possibilities: initial or response. The first group was developed to identify messages for the presence of an initial claim—a seed, which began the initial line of discussion. It consisted of two coding categories: `initial argument` and `initial assertion`. The distinction between the two was based on whether the claim was accompanied by reasoning. Messages which provided reasoned or non-reasoned claims that began an initial line of discussion and were not a response to another message’s claim or argument were coded as initial argument or initial assertion accordingly. It should be noted that this group was reserved solely for the first seed within a thread. Any additional seeds in the thread,

---

\(^{49}\) See Appendix 1 for a detailed overview of the coding phases.

\(^{50}\) Note one exception here; the unit of analysis for the category `neglected` was again the argument.

\(^{51}\) See Appendix 2 for an overview of the coding phases and categories. Additionally, examples of the categories are presented throughout the three proceeding results chapters.
which began a new line of discussion were coded as one of the two counter catego-
ries discussed below.

The second group, response, was divided into three sets of categories: it was
developed as a means of identifying the different types of reasoned, non-reasoned,
and non-claim replies—different types of interaction. A message was regarded as a
response if it directly or indirectly referred to another message. It is important to
note that these sub-groups are not mutually exclusive, and as such, a single message
may be coded multiple times under one or more of the possibilities.

The first set denoted those messages, which provided reasoned claims: counter,
rebuttal, refute, and affirmation. The main distinction between these argument types is
the relationship they share with each other. First, a message that provided a rea-
soned claim in which an alternative claim was proposed that did not directly contra-
dict or challenge a competing claim or argument was coded as a counter. Second, a
message that provided a reasoned claim, which directly contradicted or challenged a
competing claim or argument was coded as a rebuttal. Unlike a counter, a rebuttal
directly contradicts or challenges an oppositional claim or argument. Third, a mes-
sage that provided a reasoned claim, which directly defended an initial argument,
initial assertion, counter, counter assertion, non-reasoned affirmation or affirmation
against a corresponding rebuttal or non-reasoned rebuttal was coded as a refute. A
refute is a defensive response to a rebuttal. Messages that provided direct or indirect
reasoned support in favor of another particip-ant’s claim were coded as affirmations.
Finally, the second set of responses here (non-reasoned claims) were divided into
similar categories (counter assertions, non-reasoned rebuttals, non-reasoned refutes,
non-reasoned affirmations) as reasoned responses.

The final set of responses identified non-claim replies. It consisted of two cod-
ing categories: commissive and expressive. Messages that assented, conceded (partial
assent), or agreed-to-disagree with/to another participant’s claim or argument were
coded as a commissive. Messages were coded as an expressive response if they
conveyed a participant’s feeling or attitude towards him-/herself, another participant,
or some state of affairs. Expressive responses were divided into three groups: humor,
emotional comments, and acknowledgements. Humor represents complex emotional
speech acts, which excite and amuse, for instance, the use of jokes, wisecracks, and
irony. Emotional comments are speech acts that express an emotion or attitude.
Acknowledgments are speech acts that acknowledge the presence, departure, or
conversational actions of another participant, such as greeting, thanking, apologizing,
congratulating, and complementing.

52 When the content of a message matched the content of another, it was a response.
53 This was employed in order to provide a more comprehensive account of the progression and
interaction of arguments.
54 This includes an initial argument, initial assertion, counter, counter assertion, refute, non-reasoned
refute, non-reasoned affirmation, or affirmation.
55 The distinction between commissives and non-reasoned affirmations is that commissives represent
convergence between opposing claims, while the latter does not.
Phase two. During the second phase, messages containing reasoned claims were coded in two steps. The first step, evidence type, consisted of four coding categories: fact/source, comparison, experience, and example. First, fact/source identified arguments, which supported their claims by providing a fact or source as evidence. Second, an argument that supported its claim by using an analogy or making a comparison in general was coded as a comparison. Third, the category example identified an argument, which supported its claim by providing an anecdotal example (real-life, fictional, or hypothetical). Finally, an argument where a personal experience was used to support its claim was coded as an experience. It is important to note that these categories are not mutually exclusive. A single argument may use multiple types of evidence.

The second step, argument style, consisted of the coding category reflexive argument. During this step, a message or series of messages by an individual were coded as reflexive argument if they provided: (a) a reasoned claim in the form of an initial or counter argument; (b) evidence to support that argument; (c) reasoned responsiveness to challenges by providing rebuttals and refutes; (d) and evidence in support of a challenge or defense against one.

Phase three. During the final phase, all messages were coded for communicative empathy, discursive equality, discursive freedom, and sincerity. First, messages suggesting that the author had imagined his- or herself in another participant’s position, either cognitively or emotionally, were coded as empathetic exchange. Second, discursive equality contained two categories, which were degrading and neglected. A message that degraded—to lower in character, quality, esteem, or rank—another participant and/or participant’s argument, statement, or opinion in general was coded as degrading. A message coded as an initial argument or counter, which was silently neglected by the other participants within a thread—lacked a reciprocal exchange—was coded as neglected. Third, discursive freedom consisted of curbing: messages that attempted to suppress, restrict, or prevent another participant’s argument or opinion. Finally, messages that questioned the sincerity/truthfulness of another participant’s person, argument, or opinion were coded as questionable sincerity.

4.4.3 Indicators of deliberation: Assessing the quality of debate

The coding categories discussed above introduce a number of different elements of political talk. The question now is how does one determine whether a discussion forum satisfies the normative conditions of deliberation? In the paragraphs that

---

56 The other type of arguments were not included here because they represent responses.

57 Counters off the topic of discussion were not included.
follow, an operationalization of the normative conditions is provided. In particular, the empirical indicators of deliberation are discussed.

One of the difficulties with both the theoretical and empirical literature on the public sphere and deliberation is that no one explicitly defines what is high or low quality or, more importantly, what specifically satisfies the normative conditions of deliberation. For example, does a forum where 50% of the claims are reasoned satisfy the normative condition of rationality? Or does a forum, where 50% of the postings represent replies satisfy the condition of reciprocity? Most of the literature is vague when it comes to defining what is meant by high and low quality, and yet, we frequently read about this forum maintaining a high level and that forum maintaining a low level (see e.g. Strandberg, 2008; Winkler, 2005; Wright & Street, 2007). There simply have been no attempts, to my knowledge, by researchers to define precisely what they mean by such statements; what are the cut-off points, the requirements for satisfying the conditions of deliberation.

Thus, not only are the empirical indicators of deliberation identified in the following section, an attempt is made at providing an explicit account, when applicable (reciprocity and reflexivity), on what satisfies these conditions. In some cases, it is simply too arbitrary to set cut-offs as to what satisfies the condition in question. This has partly to do with the nature of some of the conditions, such as empathy, and partly due to the limitations of a textual analysis. That said, when these cut-offs are not explicit, the judgments made in the results chapters of this dissertation along with the comparative nature of the design will provide at the very least future researchers insight into developing their own cut-offs. By stating these value judgments openly and clearly, either here or in the discussion of the results, the normative framework of the empirical analysis is made transparent for agreement or disagreement by the reader. Thus, this attempt is explorative and is in no way comprehensive, but rather represents a first step for future research to build upon. However, the comparative nature of the analysis does allow us to say that one forum is e.g. higher or lower than the other for the conditions.

Rational-critical debate
The process of achieving mutual understanding is comprised of six components: rational-critical debate, coherence, continuity, and three dispositional requirements: reciprocity, reflexivity, and empathy. The first component, rational-critical debate, requires that participants provide reasoned claims, which they critically reflect upon during the course of a discussion. The literature on deliberation and the public sphere does maintain that the exchange of claims represents the guiding communicative form of deliberation. Consequently, the assessment here was achieved by

---

58 Only those coding categories that were used to operationalize the normative conditions of deliberation are discussed here. Expressives are discussed in the next section.

59 However, it is not the only relevant form.
determining the total number of postings coded as reasoned and non-reasoned claims in relation to the posting total.

It is important that the exchange of claims maintain a sufficient level of rationality and critical reflection. Rationality was assessed by calculating the number of reasoned claims (initial, counter, rebuttal, refute, and affirmation arguments) in relation to the total number of claims made (non-reasoned plus reasoned claims). While critical reflection was assessed by first determining the level of disagreement (the number of messages coded as rebuttals, non-reasoned rebuttals, refutes, and non-reasoned refutes). However, disagreeing is not always accompanied by reflection. The level of rebuttals and refutes, on the other hand, does suggest its presence because they not only include statements of disagreement, but also provide reasons in support of those statements, indicating a degree of critical reflection. Thus, by calculating the number of arguments coded as rebuttals and refutes in relation to the total number of reasoned claims made, the level of critical reflection was assessed.

Coherence
Coherence was assessed by determining the consistency of the messages within each thread. Ideally, participants should stick to the topic until mutual understanding and/or some form of agreement is achieved. Thus, the messages within each thread were first analyzed and then categorized into lines of discussion based on the issues discussed. The level of coherence was determined by assessing the number of topic changes and the relevance of such changes. The latter point is particularly important. Often discussions diverge from the original issue, for example, due to points of clarification or new issues being discovered, which are relevant to the discussion. Consequently, these types of divergences are indirectly related to the original issue and not treated as disturbances.

Continuity
Continuity requires that debate continues until understanding or some form of agreement is achieved as opposed to withdrawing from the discussion. It was first assessed by determining the level of extended debate within each thread. The level of extended debate refers to the frequency of continued interaction between participants via counters, rebuttals, and refutes. If there are extended interactions between participants in the form of rational-critical debate, then the opportunity to reach a deeper level of understanding is increased.

Lines of discussion within each thread, which were not off the topic, were coded for extended interaction via the presence of at least one strong-string. A strong-
String refers to a minimum of three argument interaction, ideally in the form of a counter-rebuttal-refute exchange. Here, an initial or counter argument is provided, which is challenged by a corresponding rebuttal, followed by a defense of that claim via a refute. If a line of discussion contained at least one strong-string, then those messages and any additional messages, which contained a claimed response (both reasoned and non-reasoned) involved in the exchange, were coded as extended debate. By calculating the total number of strong-string claims in relation to the claim total, the level of continuity was assessed.

Continuity was also addressed by determining the level of commissives. As mentioned above, continuity requires convergence as opposed to withdrawal by participants. Thus, threads were coded for acts of convergence, commissive speech acts. These represent moments during the course of a debate when a participant posts an assent, partial assent, or agree-to-disagree statement in response to another participant’s argument or position. Ideally, a line of discussion should end in some form of convergence. Consequently, continuity was assessed by determining the level these acts in relationship to the lines of discussion within a thread.

Reciprocity
In the past, net-based public sphere researchers have often measured reciprocity by determining the percentage of postings coded as replies–reply percentage indicator. The percentage of messages coded as a reply within a forum or sample of threads is calculated and used to determine the level of reciprocity. This approach focuses on measuring individual acts of reciprocity, reciprocity at a participant-to-participant level. Such an approach, however, neglects the social structure of a discussion thread; it neglects the network of messages, which connects the participants. In order to illustrate this point, a thread from the Wife Swap forum is presented below.

In Figure 4.1, the replies between participants within a discussion thread consisting of 18 participants with 23 postings were plotted. Each node (1-18) represented a participant. The size of the nodes signified the number of messages posted by each participant. The lines and arrows between nodes represented the replies and the direction from which they came. The darker the arrow, the higher the traffic was in that direction. Finally, the numbers in parentheses represented the total number of replies received and sent for each participant.

If we use the reply percentage indicator on this thread, we would find that approximately 96% of the messages posted were replies. Under this account, we might conclude that this thread had a high level of reciprocity, thus satisfying the normative requirement. However, this would be misleading, particularly if we are interested in the type of reciprocity crucial to achieving understanding. Ideally, reciprocity here

---

63 This may also include any three-combination exchange involving an initial argument, affirmation, counter, rebuttal, and refute, which represents a continuation.
64 The convergence analysis was only applied to political coherent lines of discussion. Lines of discussion containing one posting were not included.
65 This did not include initial agreement, which was coded under affirmations.
could be visualized as a web within which all the participants are connected via their postings. In this example, however, we have a centralized discussion. The initial message posted by participant one consumed the attention of most other participants thereby creating a social structure that looked more like a many-to-one reciprocal exchange rather than many-to-many web of reciprocity. Consequently, even though this thread contained a high level of replies, it still had a moderately low level of reciprocity because the social structure of those replies was centralized; participants were not listening and replying to each other but rather at one other.

Knowing the percentage of replies is of course an important factor when determining the level of reciprocity, but it is insufficient on its own as this example has demonstrated. Therefore, the level of reciprocity was assessed by combining the reply percentage measurement with a degree of centralization measurement (De Nooy, Mrvar, & Batagelj, 2005). The latter measurement was employed to investigate more precisely the social structure of the discussion threads. The concept of centrality here refers to the prominence of a particular participant. The degree of centrality indicates the number of links connecting participants to a focal participant, while the centralization of a thread refers to the degree to which centrality is monopolized by any one participant(s) in the thread. The degree of centralization for each thread was measured using Pajek, a network analysis software program. The degree of centralization was calculated by dividing the variation in degree of vertices (participants) by the maximum degree variation, which is possible in a network (thread) of the same size (De Nooy, Mrvar, & Batagelj, 2005, p. 126). Each thread yielded a score on a scale of

Figure 4.1: An example of a centralized discussion thread.
one to zero, with the former representing the optimum centralized thread and the latter the optimum decentralized thread.

\[
\begin{array}{c|c|c}
\text{Repl Percentage Indicator} & \text{Low} & \text{High} \\
\hline
\text{High} & \text{Strong Decentralized Web} & \text{Strong Centralized Web} \\
\text{Low} & \text{Weak Decentralized Web} & \text{Weak Centralized Web} \\
\end{array}
\]

*Figure 4.2: The web of reciprocity matrix.*

In order to assess the forum as a whole, the dual results for each thread were plotted along a double axis matrix (see Figure 4.2). It acted as a tool for interpreting the level of reciprocity. It was broken into four quadrants labeled: strong decentralized web, strong centralized web, weak decentralized web, and weak centralized web. Threads that fell within the strong decentralized web quadrant were considered to satisfy the normative requirement of reciprocity because they embodied both a high percentage of replies and a low level of centralization. Threads that fell within the strong centralized web or weak decentralized web quadrants were considered to have a moderate level of reciprocity. Finally, threads that fell within the weak centralized web quadrant were considered to have the lowest level of reciprocity. These threads had a low level of replies, and when participants did reply, it was highly centralized.

**Reflexivity**
The level of reflexivity was assessed at two progressive stages of coding. The first stage examined the messages for their use of evidence and set the boundaries for stage two, which identified messages for the presence of reflexive arguments. During the first stage, arguments were coded for evidence use. In everyday political talk, people reason socially on a variety of issues. When they support their reasoning or challenge others, they make use of evidence, drawing on everything from personal life experiences and observations to statistical data and media reports. Using evidence to support an argument or challenge an opposing argument indicates that a partici-
Research design & methodology

A participant has taken the time to reflect the opposing position against his/her own because, in order to relate evidence to one’s own argument or an opposing argument, a participant must know and, to some extent, understand the opposing position (Kuhn, 1991). Moreover, supporting an argument using a fact/source, comparison, experience, or example as opposed to using no evidence, suggests that a participant has reflected upon the opposing position because such evidence requires a participant to contend with questions such as where to use the evidence and what relationship exists between the evidence and the claim it supports or challenges, which requires reflexivity.

During the second stage, messages were assessed for argument style, reflexive argument. It is important to understand that reflexive arguments are usually dependent upon the exchange of numerous arguments between participants in a discussion. As such, they usually occur over a series of messages via a chain of arguments by a particular participant. When a participant posted a message or series of messages, which (a) provided a reasoned initial or counter claim; (b) used evidence to support that claim; (c) was responsive to challenges by providing rebuttals and refutes; (d) and provided evidence in support of that defense or challenge, they were assessed as satisfying the normative component of reflexivity. By comparing the number of reflexive arguments to the total number of arguments, the level of reflexivity within the sample was determined and assessed.

Empathy
Empathy is often conceptualized cognitively (mental perspective taking) and emotionally (vicariously sharing emotions). Putting yourself in another position and trying to understand matters from that person’s perspective cognitively and/or emotionally is important to deliberation. However, since deliberation is a social process, conveying empathic considerations to another participant is a critical component. When participants do not convey their empathic thoughts and/or feelings, empathic relationships cannot emerge, thus empathy has little bearing on the social process. As such, the analysis focused on capturing those instances of communicative empathy by coding for empathetic exchange. The level of empathy was initially assessed by determining the number empathetic postings in relation to the total postings.

Discursive equality
Structural and dispositional fairness is comprised of three components. The first of these, discursive equality, requires both an equal distribution of voice and that participants respect and treat each other equally. It was analyzed by assessing the rate and distribution of voice within the forum. As Schneider (1997, p. 73) states, “Equality in the idealized state would suggest that all participants ought to contribute equally—that is, each author ought to contribute an equal number of messages”. The goal here was to measure the number of participants along with their rate of participation and their share of the postings thereby determining the concentration of participation. Forums that maintain a distribution of voice skewed towards a small
group of frequent posters were considered discursively unequal, while those that were more evenly distributed, more egalitarian were considered to have satisfied the requirement.

However, such an analysis on its own is insufficient; just because participants are speaking, it does not mean anyone is listening. The question then becomes who are they listening to—the popularity of the participants. Ideally, everyone should be equally popular; no one participant or group of participants should monopolize the receiving of messages. Therefore, in conjunction with the above approach, all threads were measured and assessed by calculating the rate and distribution of popularity (concentration of popularity). By determining both the concentration of participation and popularity, a clearer picture of the distribution of voice was achieved.

The distribution of voice tells us little about the level of substantial equality within a discussion forum. Do participants respect and recognize each other as having an equal voice? This question was addressed by coding and assessing the forum for the level of substantial equality. The analysis consisted of two coding categories: degrading and neglected. The code degrading identified those instances when participants actively degraded each other. When a participant degrades another participant’s character or argument, it not only indicates a lack of respect but also creates an atmosphere of inequality. The category neglect too identified those instances of inequality. However, it focused on those instances of passive neglect, when arguments went ignored or unnoticed wordlessly.

**Discursive freedom**

Discursive freedom, the second component of structural and dispositional fairness, requires that participants are able to share freely information, arguments, and opinions in general. The aim here was to capture and describe those instances of censorship by the participants themselves, those instances when a participant was prevented from speaking his/her opinion or argument by another participant; thus, all messages were coded for curbing.

Curbing can come in a variety of forms from the use of abusive and aggressive language to direct statements of censorship. The level of discursive freedom was assessed by calculating the percentage of postings containing acts of curbing. However, it should be pointed out that not all acts of curbing impede deliberation, and in some cases, curbing may be seen as enhancing it. Consequently, all acts of curbing were initially coded and then later assessed.

**Sincerity**

The final component of structural and dispositional fairness is sincerity. It is difficult to judge whether a participant is being honest. Moreover, such a judgment would require more than analysis of the texts. The focus here then was not on whether every participant was telling the truth, but rather, it was placed on the social act of questioning another participant’s sincerity; identifying those instances when a participant questioned or challenged the sincerity of another participant. Perceived
sincerity is a crucial component to deliberation. Even if levels of actual sincerity were high, if participants do not perceive this as such, then deliberation is placed at risk.

Thus, the analysis concentrated on gauging the level of perceived sincerity, whether participants perceived others as being sincere. It was assessed by identifying those exchanges between participants where sincerity was questioned via the category questionable sincerity. If the level of questionable sincerity were high within a forum, it would be hard to envision any constructive deliberation occurring, particularly when it came to achieving mutual understanding.

4.4.4 Expressive speech acts and political talk

One of the aims of this study was to move beyond argumentation and formal notions of deliberation. The coding scheme presented above aimed at identifying expressive speech acts, common ingredients of everyday political talk. However, this analysis on its own was insufficient because it does not address the second research sub-question, which asks: What role, if any, do expressives play within online political discussions and in relation to the normative conditions of deliberation? The aim here was not only to describe how participants actually talked politics, but also to see whether expressives tended to facilitate or impede deliberation. Consequently, the above analysis represented only the first step to addressing this question. In the paragraphs below, the additional analyses that were conducted are outlined and explained.

Humor

Overall, there were four separate in-depth readings of humor conducted. In each case, the selected material was read, re-read, and worked through. Moreover, the readings were conducted in consecutive order; in other words, the first reading for all three forums was performed, followed by the second, and so forth. During the initial coding of humor, two trends/patterns were noted, which warranted further investigation. Additionally, a separate reading was already planned earlier to examine humor in relationship to certain variables of deliberation. However, the first reading focused on providing a more comprehensive account of the type of humor used.

Though an extensive list on the various types of humor was consulted during the initial coding phase, no distinction was made during the coding process. Consequently, a separate reading aimed at identifying more precisely the types of humor used was conducted. In order to carry out this analysis, additional literature on humor was consulted. In particular, Shibles (1997) guide to identifying and classifying humor was utilized.

66 In some cases this required a re-reading of the individual coded postings, while at other times this required a reading of the humorous comments in context (a reading of the whole thread). This holds true for all expressives.
Chapter 4

The second reading focused on analyzing the social structure of humor. During the initial coding, it was noted that humor seemed to invite more humor; consequently, the aim here was to investigate this further and more precisely. The analysis was made easier by the fact that the postings had already been coded for reciprocity.

The third reading focused on analyzing the participant’s use of humor. Again, during the initial coding, patterns were noted. However, before this investigation was carried out, additional literature was consulted as a means of improving the analysis. In particular, Koller’s (1988) work on the sociology of humor proved most useful.

Finally, in order to address the second part of the research question stated above, a final in-depth reading was carried out as a means of identifying any relationship between humor and variables of deliberation. For example, do participants use humor to support their arguments, or do they use it to degrade another? Given the diversity of the conditions discussed above, the analysis began by consulting MAXQDA’s code relation browser (discussed above) as a means of assisting in identifying any initial relationships between codes, intersections between codes.

**Emotional comments**

Overall, there were three separate in-depth readings on emotional comments conducted. As above, the material was read, re-read, and worked through in consecutive order. During the initial coding phase, one trend/pattern was noted, which again warranted further investigation. Additionally, a reading was already planned to investigate emotional comments in relation to the normative conditions. The first reading, however, focused on providing a more thorough account of the type of emotions used within the three forums.

Similar to above, the initial coding phase only identified a posting that expressed emotion and not the particular emotion being expressed. Thus, the first reading aimed at identifying the emotions used within political talk. Additional literature was consulted prior to the analysis in order to come to a categorization of emotions. In particular, Shaver, Schwartz, Kirson, and O’Connor’s (2001, pp. 34-35) categorization of emotions was used, which consisted of six primary emotions (anger, sadness, joy, fear, love, and surprise). Each primary emotion consists of several secondary emotions, which consist of tertiary emotions. Emotional comments were initially coded at the secondary level of this model.

The second reading focused on the social structure of emotional comments. During the initial coding, it was noted that emotional comments often came in strings; consequently, a second reading was conducted to investigate these patterns further.

A final reading was carried out as a means of identifying any relationship between emotional comments and the conditions of deliberation. For example, do participants use emotional appeal in their arguments, or are emotional arguments ignored, neglected? Again, MAXQDA’s code relation browser was consulted initially in order to assist in identifying any initial relationships between codes.
Acknowledgements

Overall, two separate in-depth readings were carried out on the use of acknowledgments. Given that they were the least frequently used expressive, the data available were limited. Moreover, unlike the above two expressives, the MAXQDA code relation browser revealed no initial relationship between codes. That said, the first reading aimed at identifying the precise acknowledgement used and exploring their use in general throughout the discussions. For example, did participants complement, what was this directed at, and who was this directed at? A second reading was conducted to investigate any findings further.

4.4.5 Validity and reliability

To increase confidence in the data collected by the coding scheme as a whole and in the individual instruments and categories, several measures were taken. One of the primary focuses during the earlier stages of developing the coding scheme was construct validity. Given the normative focus of the study, it was crucial that the empirical indicators truly reflected the conditions in question. Consequently, much attention was paid to the operationalization of these conditions.

The coding scheme was initially developed in collaboration with an expert in net-based public sphere research. An initial exploratory study was conducted to test the coding scheme instruments. During this period, particular attention was paid to the schemes functionality and workability. The results from that study are available in Graham and Witschge (2003). After this study, modifications and adjustments were made. The new version was presented and discussed at both the Oxford Internet Institute Sumer Doctoral Program and at the ECREA Summer Doctoral School (2004a; 2004b). Multiple peer debriefings were carried out with several leading experts in the field and fellow e-democracy/government researchers. Following this feedback, several modifications and adjustments to the scheme were made. In particular, multiple instruments were combined for particular conditions, for example, combing a degree of centralization measurement with the reply percentage indicator for reciprocity. This was followed by another exploratory study on threads from a Big Brother discussion forum as a means of testing the new scheme. Afterwards, additional modifications were made, and the new version was eventually published in Graham (2008).

Due to financial constraints, the coding was conducted by one researcher. To increase confidence in the data collected, an intra-rater reliability test was conducted as a means of determining coder stability. Three and half months after the final phase of coding and analysis was completed, 10% of the postings for each of the three forums were recoded for intra-rater coder reliability.

67 I like to thank Nico Carpentier, Stephen Coleman, Lincoln Dahlberg, and Peter Dahlgren for their feedback.
Table 4.1

<table>
<thead>
<tr>
<th>Scott’s Pi</th>
<th>Number of codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>2</td>
</tr>
<tr>
<td>Greater than .8</td>
<td>8</td>
</tr>
<tr>
<td>Greater than .7</td>
<td>5</td>
</tr>
<tr>
<td>Greater than .6</td>
<td>5</td>
</tr>
<tr>
<td>Greater than .5</td>
<td>1</td>
</tr>
<tr>
<td>Greater than .4</td>
<td>2*</td>
</tr>
</tbody>
</table>

*With the exception of these two codes, which received a greater than 80% agreement score, all remaining 21 codes scored greater than 90%.

Intra-rater coder reliability was calculated in two ways: percentage of agreement and Scott’s Pi. Agreement records the percentage of instances in which the coder, on two separate occasions, observed either the presence or absence of a variable. Scott’s Pi is a statistical calculation that factors in the consideration that random chance would result in a certain percentage of identical codes. Table 4.1 shows the breakdown of intra-rater coder reliability for the codes presented above. There were initially 24 codes, one of which was eliminated after the reliability test. As is shown, all but two codes of the 23 scored greater than 90% on the percentage of agreement calculation, with 20 scoring greater than .6 with regard to Scott’s Pi. According to Neuendorf (2002), reliability above 60% agreement with Scott’s Pi greater than .3 is considered an acceptable level of reliability.

4.5 Limitations

There were several limitations of this study worth noting. First, the research design was restricted to the communicative practices of participants. As discussed in Chapter 3, research has suggested that the structure of the forum plays an important role during the process of deliberation. The forum layout, the rules and guidelines, the role of the moderator, and the managing of the forum in general may influence certain variables of deliberation. Consequently, by excluding these from the analysis, there are limitations placed on the above design, particularly regarding variables such as discursive freedom and equality. Furthermore, the design was restricted to an analysis of the text, to the postings. As argued in Chapter 3, there are limitations to

---

68 Given the diversity in the frequency of codes, Scott’s Pi was chosen as the most appropriate test because it adjusts for the frequency with which categories may be used, the degree to which agreement would be expected by chance (Reinard, 2006, p. 126).
69 Second order codes were not included such as strong-strings and reflexive arguments.
70 However, Neuendorf is referring to inter-coder reliability here. That said, the standards were still used because they seemed reasonably acceptable for an intra-rater test.
71 The rules and guidelines and any literature made available by the forums on the role of their moderators and managing practices were initially consulted for any blatant problems.
Research design & methodology

what can be detected concerning variables like reflexivity. Reflexivity is largely an internal process of understanding—reflecting another participant’s claim against one's own. Consequently, by limiting the analysis to what was being posted, the actual processes that take place within the minds of participants were neglected. Although, as was demonstrated above, reflexivity to a certain extent can be deduced from the arguments provided by participants, ideally such an approach would be complemented by interviews as a means of providing a more comprehensive indicator.

Another possible limitation of the design was its complexity; the design was multifaceted, complex, and extensive at times. That said, owing to the number, diversity, and complexity of the variables of deliberation, a thorough and comprehensive operationalization of the normative conditions was required; one that would allow the creation of indicators, which actually reflected the normative conditions in question.

One limitation that was beyond the control of this study was the editing practices of the forums. During the analyses, there were two incidents identified where postings had been noticeably modified or removed within two of the three forums. These modifications and deletions were detected by chance via the practice of participants to include the message they were responding to in their postings. Consequently, postings that had been deleted and modified were detected here. The webmasters of these sites were contacted to see if it was possible to obtain the original postings, but in both cases, the attempt was unsuccessful. Consequently, it is unclear how often these types of events occurred and what consequences they had on the analyses.

Finally, the researcher in this study had limited contextual awareness and was foreign to the political culture. The forums under investigation were all British-based and the researcher examining them was not British nor had lived in the UK prior to the analyses. Consequently, there was the chance that e.g. inside jokes, local and regional terminology, and events might have been missed or misunderstood. However, every effort was made to familiarize the researcher with the material. On occasions, British colleagues were contacted to inquire about, for example, slang words, possible humorous comments, or any difficult statements in general.

---

72 These postings were integrated back into the threads when enough information was available with one exception, when the posting occurred towards the beginning of the thread. If most of the participants were responding to the modified version, then it was included.