

DECODING DELIBERATION ONLINE

Jennifer Stromer-Galley

Paper presented at the Second Conference on Online Deliberation, May 2005, Palo Alto, California.

Jennifer Stromer-Galley
Department of Communication, SS 340
University at Albany, SUNY
Albany, NY 12222
jstromer@albany.edu
518-442-4873

The literature on deliberation has grown at a rapid pace in the past fifteen years. A number of creative and influential experiments have found that deliberation affects political knowledge (Fishkin, 1991, 1995), consideredness of opinion (Luskin, Fishkin, & Iyengar, 2004), and attitude and opinion change (Cappella, Price, & Nir, 2002; Gastil & Dillard, 1999; Price, Cappella, & Nir, 2002), to name a few effects. A comprehensive review of the deliberation literature by Ryfe (2005) indicates, however, a large hole in the research concerning what actually transpires when people “deliberate.” Few studies attempt to measure how deliberative a group’s interaction is, or how the process of deliberation within a group works.

This paper sketches elements that might undergird normatively “good” deliberations. Based on empirical research as well as the theoretical literature, several “ingredients” of deliberation have been identified. These elements have been operationalized and a coding scheme generated to help track what groups are actually saying and doing when they deliberate.

The aim of this paper is modest. Before actually applying this coding scheme to a set of deliberative exchanges to help test research questions and hypotheses about the quality of group deliberation, the role of specific discursive acts, such as disagreement, in the process of deliberation, and the ways groups move to consensus, the coding scheme needs to be pilot tested. This paper, then, explains the ingredients of deliberation culled from the literature, and the modest testing of the coding scheme using publicly available political discussions that occurred online. The results of the pilot test suggest that a few elements of the coding appear to be working quite well. Others, however, need additional tailoring.

Studying the Black Box

Several studies and experimental interventions attempt to create opportunities for political discussion and deliberation to occur. One of the most publicized is James Fishkin’s

deliberative poll, which brings citizens together, provides them detailed briefing packets on the issue under consideration, and then surveys the participants both before and after the discussions. Fishkin's (1991, 1995) research has found significant learning on the political issue and substantive changes in opinion following these deliberations.

This model of experiment is fairly common, whereby citizens are recruited, either randomly or by convenience, to participate in a deliberation, and are surveyed on their knowledge and opinion levels before and after the deliberation to determine what effects on knowledge and opinion the deliberation has.

Less known is what happens during the course of the group discussion itself. Few studies look deeply and in a systematic way at the discussions to determine what citizens are saying to each other. Generally, it is assumed that what the citizens are doing is deliberation, yet rarely are measures taken to assess the quality of that deliberation. What we do not have a clear sense of is how citizen discussion works, and how external forces, such as moderation, demographic and ideological composition of the group members, and channel (in person versus online), for example, affects the group process.

One research study suggests that specific effects can be traced to what transpires in the deliberation itself. Lemus et al (2004) analyzed long-collaborating computer-mediated work groups that were tasked with solving a problem. Those groups that developed the most arguments for a solution, with the most people contributing to those arguments, were more likely to see their argument become the basis for the solution to the problem. The number of reasons also was a variable predicting which solution was advanced. Thus, it seems important to begin measuring what occurs during deliberation.

It should be noted that the assumption in the research described in this paper is that although the individual is an important component of the deliberation, it is equally important to recognize that deliberation is a product of social interaction, and is structured and created by a group of people. As such, communication between people is the essential mechanism through which a group of people create a reality (Burke, 1945). Thus, although individual psychology plays a role, equally important is the process and product that a group of people create in deliberation.

Yahoo! Political Chat

The ultimate aim of this coding scheme is to apply it to experimentally-generated dialogues on political topics, comparing groups who use different media through which to interact. Yet, before the experimental work can begin, it seems important to work out and fine tune a coding scheme capable of discerning differences that might occur between groups during political dialogue.

For the purpose of testing the coding scheme, Yahoo!'s Politics chat room discussion was used as the data source for this pilot project. Aside from being a convenient forum on which to test the coding, online chat is generally considered to not be very good political dialogue. The assumption is that online chat is filled with *ad hominem* talk and fragmented ranting, and with people who do not wish to talk but do wish to shout at other people who happen to be there. Yet, prior observations of online chat as part of a prior research study (Stromer-Galley, 2002), suggest that there are moments of rich political discussion. If there is variability, then, in the *prima facie* quality of political talk in Yahoo Chat, it serves as a useful forum on which to test the coding scheme.

The chat sample to be coded was drawn over six days from January 23 to January 28, 2004. Two hours per day were recorded. Recording was done during the evening hours.¹ In total, 12 hours of chat were recorded over that period. All 12 hours were coded following the coding scheme described in the following section.

The Ingredients of Deliberation

In this section, several “ingredients” for deliberation and dialogue are discussed. These ingredients form the basis for the coding scheme to be pilot tested. Literature was culled from several areas of research including deliberation, political talk, small group, and online interaction. As well, the theoretical and normative literature on political talk was also consulted. The aim was to identify shared or agreed upon characteristics that comprise political discussion and deliberation.

There are a handful of studies that operationalize deliberation for content analysis. Interestingly, these studies have primarily coded discussions that occur through the Internet. To date, I have found only a few studies that systematically code face-to-face group political deliberation (see, for example, Dutwin, 2002). At this point in the literature exploration, it appears that there are eight “ingredients” that are essential for good political dialogue: reasoned opinion expression, references to external sources, lack of ad hominem attack, clash/disagreement/opposing perspectives, equal levels of participation, coherence, reflexivity, interactivity, and narrativity. Each is taken in turn and described in some detail.

Reasoned opinion expression

Rational argument is one of the cornerstones of deliberative and political theory. Many of the studies that analyze political deliberation, experimental or organic, reference the Habermasian public sphere, which requires rational argument among acquainted and

unacquainted others in order for public opinion to form and to serve as a check on formal government. Rational argument requires individuals to advance positions that can be defended against critique. In *The Theory of Communicative Action*, Habermas (1984) argues that an expression is rational if the utterance provides evidence that can be observably confirmed or denied empirically or appeals to a shared normative ground. Thus, in a deliberative context, a reasoned argument would be one in which assertions are grounded in empirically verifiable evidence or in shared understanding of moral or normative behavior.

In studies of dialogue, reasoned argument generally is defined as opinion claims supported with evidence for those claims. Dahlberg (2001), for example, in his model of ideal deliberation requires “exchange and critique of reasoned moral-practical validity claims” (p. 623). In other words, “deliberation involve[s] engaging in reciprocal critique of normative positions that are provided with reasons rather than simply asserted” (p. 623).

Similarly, Graham and Witschge (2003), advocate for a coding scheme that includes rationality. They coded a set of political conversations on UK Online, a publicly available, asynchronous, threaded, group discussion. Each message was coded first on type of message in the sequence (an initial message, a response to a prior message, or an irrelevant message not part of the discussion topic). Second, each message was coded for whether it was a rational or an irrational message. A rational message was one that provided justification or a reason for a “validity claim” (p. 182). A third code further tracked messages for the type of argument the message provided.

In keeping with their work, as well as that of Cappella, Price, and Nir (2002), one category of code for this pilot project is reasoned opinion. Messages that provide an expression of opinion or provide an assertion is coded on whether there is only an opinion or an assertion, or

if that opinion or assertion is supported with some reason, elaboration, or justification for the remark. For example, an utterance that opines “Bill Clinton was an excellent president” would be coded as an opinion without a reason. A message that opines “Bill Clinton was an excellent president. His stewardship over the economy allowed many of us to experience prosperous times” would be coded as an opinion with a reason. In addition, after several rounds of practice coding, it became clear that a third category, “facts,” needed to be included in the coding scheme. “Facts” are coded as stand-alone statements of reality that do not appear to have any opinions associated with them. Although it might seem odd that dialogue on political topics would produce stand-alone facts, a number of instances were found where that was the case. Sometimes, the fact came to serve eventually in support of an opinion to create reasoned opinion, but not in all cases.

Sourcing

The coding conducted of online dialogue of prior research generally includes a code for reference to an external source (Davis, 1999; Graham & Witschge, 2003; Hill & Hughes, 1998). Another way to measure whether deliberation is rational is whether references are made to empirically verifiable sources of information, such as newspapers, websites, broadcast news, or particular people who are recognized as knowledgeable or authoritative on the subject matter. For example, Davis (1999) conducted a content analysis of political threads on Usenet, which included a code for “evidence.” This included “inclusion of other materials” defined as “references to, quotations from, or discussion of materials such as specific books, articles, government reports, speeches, etc.” (p. 162).

This project, then, includes a code for sourcing. A message is coded as providing a source if it makes explicit reference to books, articles, government reports, speeches, websites, television broadcasts, experts, or personal experience.

Flaming

Ad hominem attack, “flaming” in internet jargon, is a source of concern, particularly of political talk that occurs through the Internet. There is a pervasive assumption that most group dialogue online involves a high level of personal attack, although research by Hill and Hughes (1998), for example, concludes that there is not as much attack in online discussion groups as popular perception would suggest. Personal attack, particularly a high degree of it, indicates dialogue that is irrational, and exhibits our worst fears about political talk. Exchanges of the sort, “well, you’re an idiot if you think that,” do not embody the high-minded ideals of political discussion; instead, they look like teenagers having a fight (White, 1997).

Ad hominem attack, then, is another category of interest in this study. A message is categorized as exhibiting *ad hominem* attack when the message attacks the speaker rather than the claim, evidence, opinion, and the like.

Disagreement

Disagreement is an important marker of political discussion. Disagreement is the signifier that there is a problem in need of a solution, a conflict in need of consideration and resolution. It also is a sign that there are participants in the dialogue with distinct views on a particular issue. This difference suggests heterogeneity of perspectives.

The question of whether there is disagreement in a dialogue matters for three reasons. The first is that there is a pervasive sense that people who engage in political dialogue online in publicly accessible spaces are engaging with like-minded others, creating issue enclaves

(Doheny-Farina, 1996). So, people in favor of legalized abortion converge with others who share the same belief. Those who oppose legalized abortion share their concerns. These like-minded groups lead to a fragmentation of the public sphere, as people choose only to talk with people like themselves and to be exposed to people whose perspectives are similar to their own (Sunstein, 2001). The second is that people who share similar perspectives are more likely to polarize in their beliefs; that is, they are more likely to develop more extreme attitudes as a result of their interactions with like-minded others (Sunstein, 2003). When there are participants in the dialogue with alternative perspectives, this can mitigate the polarization effect (Sunstein, 2003). Third, people who are opposed to disagreement, to views that are different from their own, are more likely to have their own views clarified and cemented (Cappella et al., 2002).

Thus, measuring whether disagreement is occurring in a political discussion can serve as an indicator of the level of heterogeneity of the participants and can provide some indication of whether participants are experiencing the benefits found in Cappella et al.'s study (2002).

For this coding project, then, disagreement and agreement is coded. A message that expresses explicit disagreement or agreement (with statements such as, "I agree," "I disagree," "you're wrong") are noted. Attempts were made in practice sessions of coding to code code for implicit agreement and disagreement. It became clear that such coding was too difficult to discern in the organic, non-topic specific arena of Yahoo! Political chat. In a dialogue with a clear, pre-determined topic, like that found in an experimental setting, it would be possible to code for implicit agreement and disagreement. For this exploratory project, however, it had to be abandoned for the more concrete coding for explicit agreement and disagreement.

Narrativity

In contrast with literature on argumentation and citizen deliberation, Dutwin (2002) argues that most citizens fail to meet idealized notions of argument and deliberation. Instead, he argues, citizens reason and offer evidence based on personal narratives. He writes, “with regard to what deliberative participants say and how they say it, the literature on argumentation, conversation, and everyday reasoning strongly suggests that the principle of reasoned and rational argument, especially on public policy issues, will not likely be a common feature of deliberative dialogue.” Instead, he predicts that “communal narrative” and “individual experience” will mark the discourse (p. 49).

From direct observation of Yahoo!’s political chat, there are virtually no personal stories, individual experiences, or communal narratives that serve as arguments. Therefore, that element is not included in this coding project. However, it is conceivable that other forums for deliberation (on or offline) might exhibit this form of discourse. As a result, this ingredient is noted here.

Equality

Along with reasoned opinion expression, equality is theorized as an essential element of healthy deliberation. In deliberation, the thinking goes, each participant must be able to participate on equal footing with every other participant (Dahlberg, 2001; Graham & Witschge, 2003; Habermas, 1984). No participant should dominate the conversation or silence others. This, too, is of concern in online dialogue, because research of Usenet discussions suggests that although a fairly high number of people seem to follow the discussion, a small minority actually contribute to the discussion (Davis, 1999).

Equality is measured by counting the number of unique people who participate in a conversation, the frequency of participation, and the number of people who “enter a room” but who do not contribute to the public discussion. The volume of messages by participants also can be noted to see if there is domination over others in the discussion.²

Interactivity

Another measure of whether a dialogue meets the ingredients for good dialogue is genuine interaction between discussants. Similar to concerns of equality and coherence, it’s important to note whether participants are actually engaging each other, or if they are simply engaging in monologues with themselves to an audience that fails to respond. Graham and Witscghe (2003) in their coding refer to this as “reciprocity,” where people take turns speaking and respond to the claims of others. As a result, who is responding to whom is tracked, and it is also noted if subsequent messages are responses to prior messages.

Reflexivity

One of the principles identified in the deliberation scholarship is a concern about reflexivity. At issue is a concern that individuals in a dialogue not only turn-take freely and participate equally in the conversation, but also demonstrate a genuine attempt to understand other perspectives.

It stands to reason that the best way to measure this is a survey of or interviews with the participants, because direct observation of reflection and reflexivity, especially of an online discussion, seems difficult without asking respondents. Because this ingredient appears to be better measured through talking with participants rather than observing their discourse, it is not included in the coding scheme.

Perhaps One More: Speech Act or Response Type

The coding of deliberation should, at least, account for the eight items noted where applicable. In the case of the present exploratory study, five of the eight are accounted for: reasoned argument, disagreement, *ad hominem* attack, interactivity, and sourcing. It also seems useful to pay attention to one other element not mentioned in prior deliberation research. That element is noting what an utterance is *doing*.

Speech act theory reminds us that language is action. It shapes how we come to understand reality. It reminds us that different kinds of talk do different kinds of things. Searle (1969, 1979) offered a typology of speech acts that organizes utterances into what he described as a direction of fit between the world and the words. Some utterances make the words match the world. Other utterances make the world match the words. Others can do both or neither. In the end he identified five types: assertive, commissive, directive, performative, and expressive. Some scholars have used this speech act typology to analyze online interactions, primarily of business oriented or work group interaction (Twitchell, Adkins, Nunamaker, et al, 2004).

For this project, each utterance in a discussion is coded using the five types of illocutionary acts developed by Searle (1969, 1979), elaborated on by Vanderveken (1990), and further modified by Cooren (1999). Cooren's conceptualization of the five types were adopted. They include: the *assertive point*, which states as actual a state of affairs ("Nuclear power is a better alternative fuel source than solar power"); the *expressive point*, which expresses a psychological state of the speaker and gives something from the speaker to a hearer, and places some value (positive or negative) on what a speaker is doing or saying (approval, sanctioning, encouragement, removal of blame); the *directive point*, which aims for the speaker to get the hearer to do some future action ("Stop arguing."); the *commissive point*, which commits the

speaker to do some future action (“I’ll be back in a second.”); and the *performative point*, which attempts to bring into existence some state of affairs by representing the speaker as performing the action (“I now declare you husband and wife.” “I pronounce you guilty as charged.”). This last category was modified for the purposes of this coding project to also include utterances unique to the online environment, such as role playing, where a speaker declares that he or she is doing something that cannot be physically done in the online environment (“I’m doing the tango;” “I sit in the corner and pout.”). This also was used to include nonverbal actions, such as “LOL” (laugh out loud).

Testing the Coding Scheme

The remainder of this paper focuses on identifying whether and how the coding scheme is working. The analysis suggests that the coding scheme is off to a good start, especially the interactivity and reasoned argument codes, but more work is needed.

Coding Deliberation

It should be noted that the content analysis was conducted following Neuendorf’s (2002) guidelines on content analysis. The author and a coder trained together to achieve an acceptable Cohen’s Alpha intercoder agreement. On all items, the coders achieve an Alpha level of .8 or higher. Coding was done at the level of the utterance. Put another way, each participant’s unique entry into the discussion was coded. This posed complications, because some participants had much to say but strung it out over several utterances; others, however, would type their long messages in their entirety before entering it into the discussion. This created a disparity in volume. The solution was to track who was continuing themselves so as to keep a record of that phenomenon.

Putting the Coding to the Test

Interactivity

The interactivity code, which measures whether people are interacting and responding to others or are engaging in monologue with themselves (currently when a new topic is introduced the coding scheme rule is that a 0 is given for the non-interactive) (Cohen's Kappa = .87). As Table 1 indicates, the coding scheme differentiates between discussions that are interactive from those that are not. The discussions on January 23, January 25, and January 27 are noteworthy for having proportionally less interactive exchange than the other days. The discussion on January 23, for example, was entirely a social discussion with only a few people participating. Many people came and went without engaging for any prolonged period with others. The discussion on January 25 included two participants who primarily were in monologues with themselves. In effect, both gave long speeches. One elaborated on the failures of George Bush as president and that Richard Clarke should be the next president. The other, over a period of two hours, talked almost exclusively to himself about the atrocities and failures of the war in Iraq. By contrast, the discussion on the 28th was a highly interactive discussion. Several people participated in discussing their views of the parliamentary system, the civility of politics outside the United States, and liberalism inside and outside the U.S.

[Insert Table 1 about here]

The level of interaction in a given discussion can be seen even more clearly when the many threads that are intertwined in pseudo-synchronous chat are disentangled. Five distinct threads were pulled out of the overall discussion for further analysis. The five threads included three pulled out of the discussion on the 24th. These three threads included several rounds of ad hominem attack (labeled "flamefest"), a long 500 utterance exchange about homosexuality and

whether it was a sin (labeled “homosexuality”), and an equally long exchange about the war in Iraq (labeled “war”). The other two discussions include the monologues from the 25th (labeled “monologue”) and a discussion on the 28th about malpractice insurance and whether something should be done to curtail lawsuits. This discussion on the 28th included a relatively large number of utterances that included references to an external source, which is why it is labeled “sourcing.” Table 2 illustrates that the three discussions produce quite different amounts of interactive exchange. The monologue on the 25th unsurprisingly produces the greatest number of non-interactive exchanges. By contrast the discussion on the 28th about tort reform is a highly interactive exchange.³ It is noteworthy that the flameword from the 24th has almost as many interactive utterances as non-interactive. This personal attack involved primarily two people, and they ranted in long monologues about each other as much as they engaged each other directly in an attack.

[Insert Table 2 about here]

Reasoned Opinion Expression

The reasoned argument code measures how much opinion, reasoned opinion, and fact is uttered in the discussions. Opinions are assertions, such as “We needed to go to war in Iraq” or “George Bush has been a terrible president.” Reasoned opinions offer some additional justification, explanation, evidence, and the like, to support the opinion, such as “Iraq was a threat. We needed to go to war in Iraq.” Facts are neither opinions nor reasons, but are stand-alone statements that are verifiable, such as “The U.S. uses 10% of the world’s oil” (Cohen’s $\kappa = .78$).

Table 3 illustrates the results of the amount of reasoned argument per day. Of special note is the discussion on January 23. The two hours recorded from that evening consisted entirely of

social chat. “Social” means that the discussion was on topics other than politics. Visitors to the chat room, when there was any discussion happening, were presented with topics about football, popular television shows, and questions about where everyone lived. As a result, there are few utterances that have opinion content. On the other days, there is more opinion and reasoned opinion utterances than non-opinion utterances, suggesting that the majority of the conversation is producing opinion-driven conversation. It is noteworthy that on one day, the 25th, there is almost as many reasoned-opinion statements as pure opinion statements. The discussion on the 25th included the monologues from the two men discussed in the Interactivity section. Their monologues were heavily reasoned, with long, fairly complicated arguments. By contrast, the discussion on the 24th, which included the discussion about the war in Iraq and whether homosexuality was a sin had approximately the same percentage of opinion statements, but there were fewer reasons or explanations given for their opinions. Likewise, discussion on the 27th offered relatively few reasoned opinions, but a great deal of raw opinion. The discussion that evening produced the greatest volume of messages over the two hour period (over 4,000, compared with the average 2,000 for the other evenings), with much hyperbole attacking liberals in the United States, and sustained attack on the Clinton presidency.

[Insert Table 3 about here]

In looking more closely at specific threads, the picture of reasoned argument becomes clearer. Table 4 provides the results of the analysis by type of discussion. The “flamefest” on the 24th produces virtually no reasoned opinion. On the face of it, that might not seem intuitive, since personal attacks might seem to be grounded in opinion statements. If a person says “you’re an idiot,” that might be viewed as an opinion. However, as will be discussed later in the response type section, speech act theory, upon which much of this coding is based, would characterize

such utterances as not *asserting* but *expressing*. Opinions are assertions; however, personal attacks are not assertions, therefore they are not opinions in this coding scheme. Instead, personal attacks sanction another speaker—generally by socially ostracizing the other person. Hence, the “flamefest” has very little opinion in it. The discussion on the 25th, characterized as the monologue exhibits a healthy percentage of reasoned opinion expression. The monologues were noteworthy in the complicated arguments presented, with several utterances serving as support for an opinion. Thus, the monologues, although not offering much by way of interaction with others, nonetheless produced a healthy amount, nearly 40%, of reasoned argument. Likewise, the discussion on the 28th about tort reform, labeled “sourcing,” produces the largest amount, 46%, of reasoned argument. The discussion on homosexuality on the 24th, by contrast, edges at times into absurd claims about homosexuality being an evil in society without much proof for such assertions, which explains the reduced amount of reasons for opinions. About 5% of the discussion, included fact statements. Such statements included statistics about the mortality rate of homosexuals being higher than that of heterosexuals. It is also noteworthy that almost 60% of the utterances were of some other sort than opinion statements. The conversation included some *ad hominem* attack as well as many directives and questions, which are not coded as contributing opinion.

[Insert Table 4 about here]

Disagreement and Agreement

The agreement code was intended to track the volume of agreement and disagreement occurring in the discussion. As mentioned before, in the early stages of the coding training, the coding for agreement included coding for explicit agreement or disagreement. Such statements included phrases, such as “I agree,” “you’re right,” and the like. The coding also included

implicit agreement and disagreement, which meant paying close attention to whether subsequent utterances appeared to agree or disagree with prior utterances. After several rounds of practice, it became clear that the coders could not achieve a sufficient level of agreement on whether an utterance was implicitly agreeing or disagreeing. Therefore, the coding, for this pilot test only includes the codes of explicit agreement or disagreement. (Cohen's Kappa: .84)

[Insert Table 5 about here]

Table 5 shows the results of the agreement/disagreement code across the six evenings of discussion. In general, there was very little explicit agreement or disagreement happening in these discussions. In looking more closely at the specific threads in Table 6, again the coding suggests that there is little explicit agreement and disagreement. The discussion on the 28th, it should be noted on tort reform, provides the largest amount of disagreement at 10%. This conversation focused on a highly specific policy, unlike the discussion on homosexuality and on war. It is possible that the discussion on a specific policy generates more explicit agreement and disagreement, although that speculation requires further data for testing.

[Insert Table 6 about here]

Although the coding detects little explicit agreement and disagreement, the coders observed a great deal of agreement and disagreement in the discussion. Because the discussion on Yahoo! chat is unstructured, it is difficult to track implicit agreement and disagreement. Unlike an experimental setting or a structured discussion on a policy position, Yahoo! chat allows discussants to range far on the topics they discuss and the level of specificity and focus they bring to any particular topic. Perhaps, in coding deliberation generated through an experiment, it will be possible to code for implicit as well as explicit agreement and

disagreement, and achieve a sufficient level of intercoder agreement to make such a code reliable.

Sourcing

The source code was intended to track whether discussants supported their arguments with websites, books, television, the news, other people, or personal experience. (Cohen's Kappa: .85) Similar to the agreement and disagreement code, Table 7 suggests that participants did not make many explicit references to external sources to help support and frame their arguments. When looking at the specific threads, as shown in Table 8, the discussion about tort reform on the 28th produces the largest percentage of sourced utterances.

[Insert Table 7 about here]

[Insert Table 8 about here]

For future coding analysis, more work needs to be done to tease out the sources people rely on in making their arguments. Gamson's (1992) work serves as a useful place to think through ways to code sources that are referenced implicitly. It is an open question whether intercoder agreement can be achieved when attempting to track implicit reliance on sources, but for the next attempt, effort will be made to that end. The narrativity "ingredient" discussed earlier also functions as a kind of reasoning or framing process, suggesting that narrative should be part of a code on sourcing rather than a separate category for analysis.

Flaming

The personal attack code tracks "flaming" or *ad hominem* attack. Utterances that included any statement that attacks the person rather than the argument was noted as being "personal attack." (Cohen's Kappa: .88). As Table 9 illustrates, overall, there is little personal attack in the two hour discussions over the course of six days. At its worst, it comprised 7% of the discussion,

on the 27th. It should be noted that none of the days analyzed had no personal attack. That is, personal attack does occur on occasion during these discussions. Yet, it is not a pervasive facet of the online discussions—at least during the hours analyzed.

[Insert Table 9 about here]

Looking more closely at individual threads, Table 10, shows that when there is a flamfest, the coding does pick it up. The discussion on the 24th was a thread that included on the face of it a large amount of personal attack. In the discussion 51% of the utterances lobbed an attack on the individual rather than on the argument. The discussion about homosexuality on the 24th, which included on the face of it some personal attack, was comprised of personal attack only 2% of the time.

[Insert Table 10 about here]

Response Type

The response type code was intended to track what utterances were doing in the discussion. The response types included assertions, expressives, directives, commissives, and performatives. (Cohen's Kappa= .87) For the pilot testing, the codes adhered to the categories established by Searle (1979) and refined by Cooren (1999). The analysis of the six evenings as shown in Table 11 suggests that the overwhelming majority of the discussion was assertions. On the 25th, nearly 75% of the discussion was constituted by assertions. At the lowest, on the 23th 51% of the discussion was assertion. The discussion on the 23rd was the social discussion, which relative to the other evenings produced the largest percentage of performative and expressive discussion. Utterances that were categorized as performative could be laughing as well as play that is unique to the online environment. Such play might take the form of “sally dances with pete” or “katy tickles dade.” Such dancing or tickling does not actually happen, but a playful

reality is created through such utterances. Expressives, in this coding, include greetings and thank yous, as well as statements that praise or sanction another discussant.

[Insert Table 11 about here]

In the analysis of the six unique threads, as shown in Table 12, the coding reveals that when the discussion is primarily ad hominem attack, then expressive utterances dominate the discussion. The discussion on homosexuality on the 24th shows the largest percentage of directive utterances. The directive category includes questions directed to others as well as commandments for others to do something. One participant in the homosexuality discussion tended to ask questions of others rather than make assertions himself, which appears to have helped elevate that group's discussion.

[Insert Table 12 about here]

The response type category is still unsatisfying. Because the categories offered by speech act theorists combine theoretically functionally-similar utterance types. From a theoretical perspective asking questions (“what time is it?”) and giving commands (“stop shouting”) are functionally the same. They both direct the listener to do something—to answer a question or to respond to the directive. Yet, questions and directives have different implications within the context of a discussion about political topics. Genuine questions that attempt to draw information out of others in the discussion shape the conversation quite differently from a command. It is possible that commands are the start of *ad hominem* exchanges or are part of *ad hominem* exchanges. Two people who disagree with each other could move to a command that one of them “shut up,” That command could then lead to personal attack. Questions, in contrast, are more likely to keep the conversation on the topic.⁴ Although it would have been better to realize

some of these issues before doing the coding, it is results like this that make pilot testing important.

A Quick Summary

Of the six evenings of discussion recorded, the prima facie “best” discussion appeared to occur on the 26th. The conversation that evening revolved around the Senate’s report on the 9/11 terrorist attacks, public opinion on abortion, and the location of the ancient city of Troy. There was also talk about the Democratic presidential primaries and the chances of Bush losing the general election. There were several people apparently committed to these discussions, and they offered their opinions, often with evidence, and occasionally even some references to sources to frame and support their positions. The fairly genial conversation began to disintegrate towards the end of the two hours when a new participant enters and begins personally attacking people and baiting them into fighting with him. The coding scheme seems to have picked up the good qualities of this discussion, such as a high level of interaction, and a good amount of reasoned argument.

By contrast, the discussion on the 24th is somewhat of a mixed bag. The homosexuality conversation is intermingled with a conversation about the war in Iraq. The homosexuality thread does not produce the same volume of reasoned argument as the discussion on the 26th, in part because the conversation over two hours appears to primarily be a subtle attempt by a homophobic person to bait a homosexual person into admitting that homosexuality is a sin. The strategy of baiting is not genuine deliberation, but a strategy intended to win the argument. The coding scheme appears to be picking up on that, although, as mentioned above, more refinement of the coding scheme is needed.

Next Steps

The pilot testing of the key elements of the coding scheme have been helpful in identifying where the coding scheme seems to be working—such as interactivity, and reasoned argument—and where some more work is needed to further refine the coding—such as disagreement, sourcing, and response type.

The coding scheme will be used in a revised form for a couple of different projects. One project will further focus on the political talk on Yahoo! chat and attempt to identify what triggers apparently good deliberations to turn undeliberative. A second project will use this coding scheme to study deliberations generated through an experiment conducted at Carnegie Mellon University. Called the Virtual Agora project, the deliberations focused on the problem of school consolidation in Pittsburgh. Over 2 sessions lasting an hour and a half each, discussants talked through four solutions to the problem of half-empty classrooms in the slowly depopulating city. The research will address questions about how groups bond and identify with each other, how they move to consensus, how they use information provided to them in advance of the discussion in the deliberations, and the role of agreement and disagreement in their deliberations. With a further improved coding scheme, it is hoped that a deeper understanding can be achieved of how groups deliberate and whether some groups do it better.

Table 1

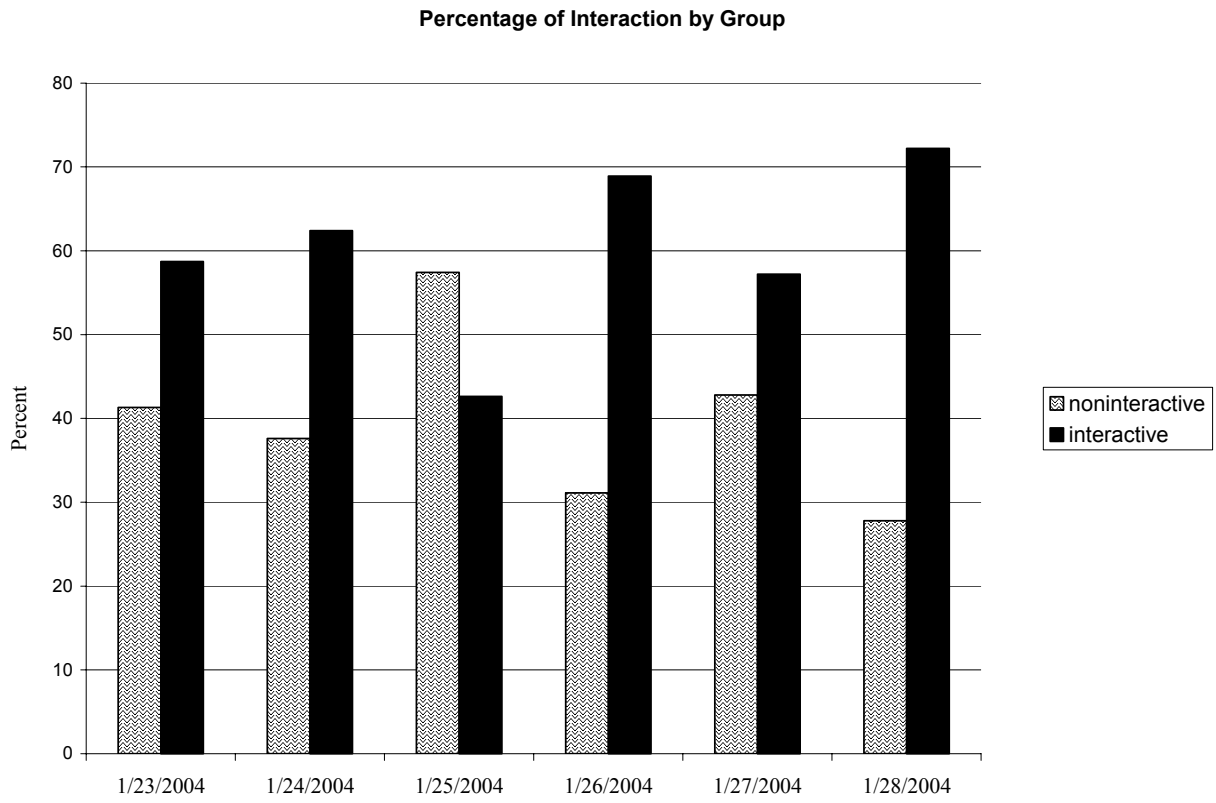


Table 2

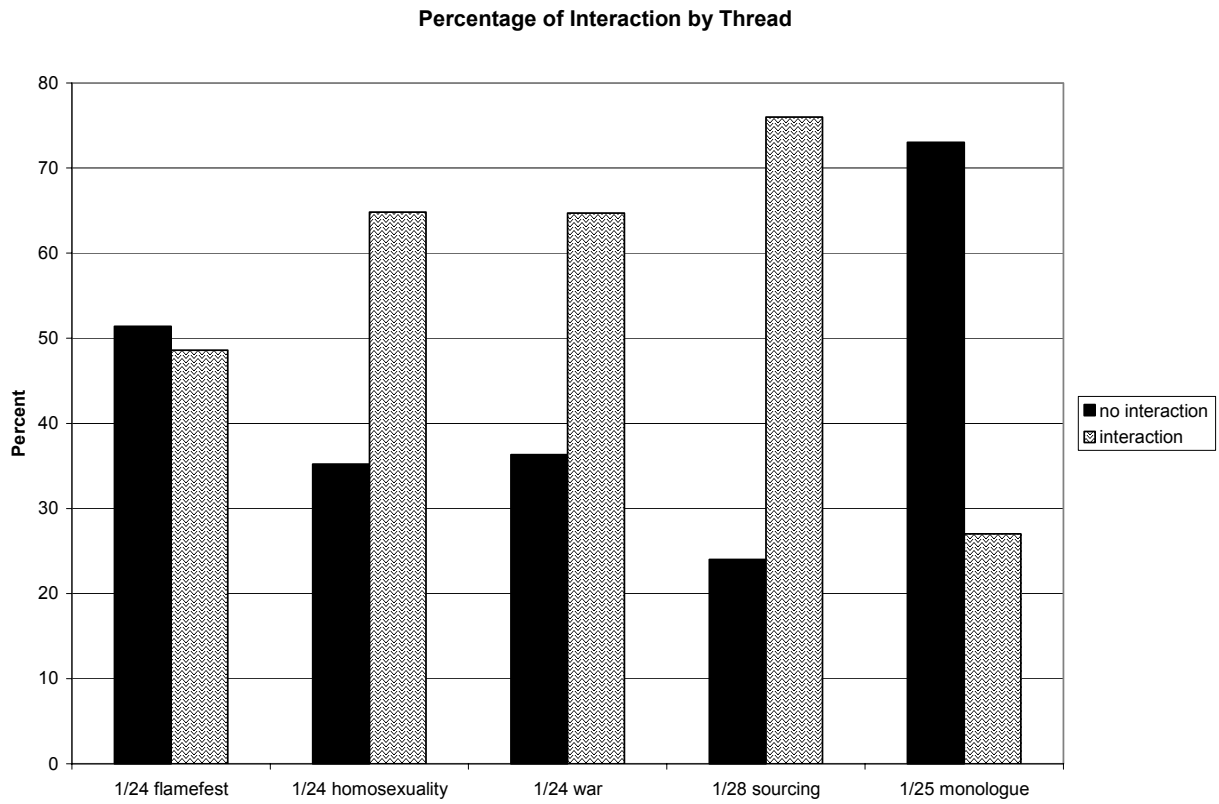


Table 3

Percentage of Reasoned Argument by Group

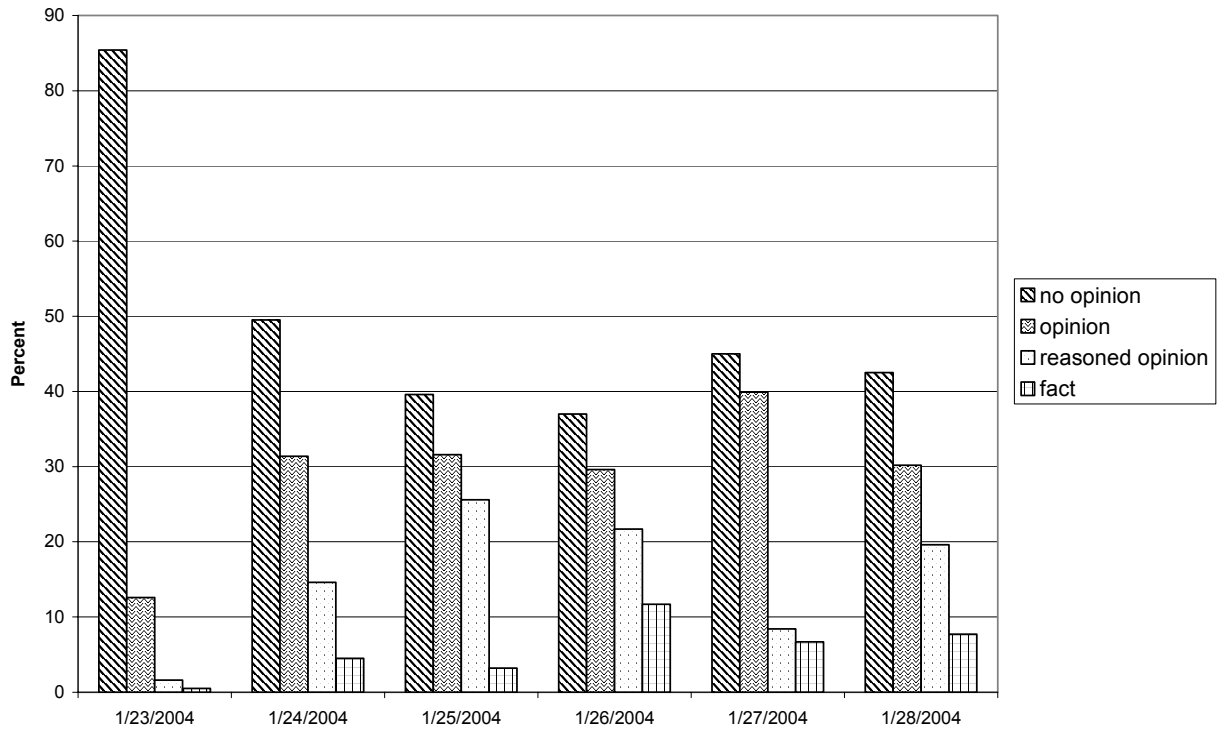


Table 4

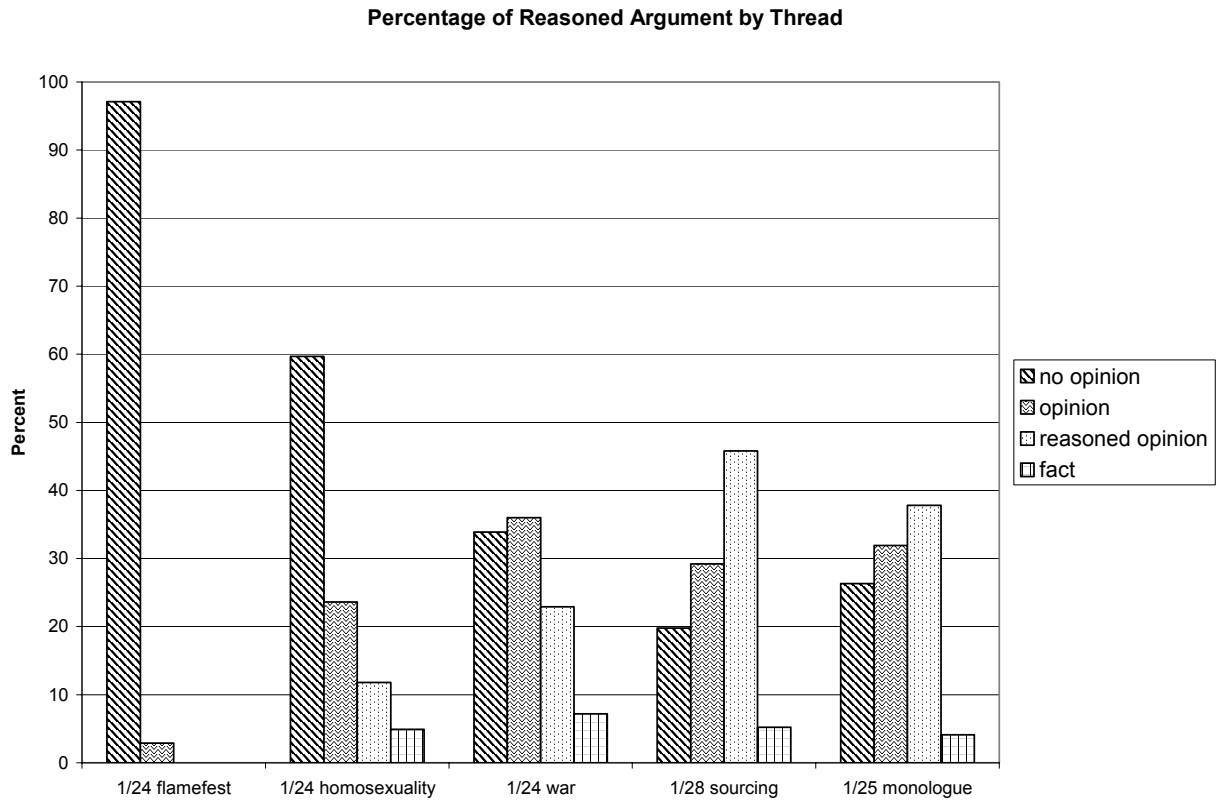


Table 5

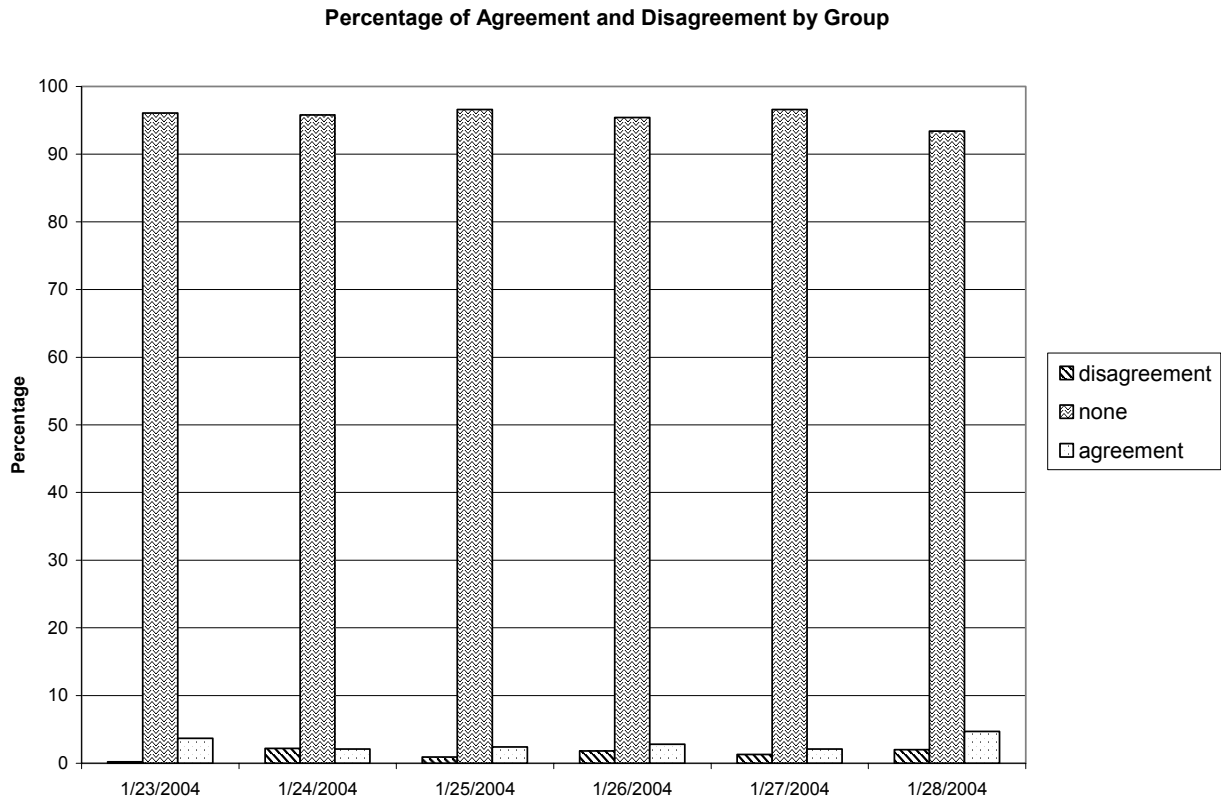


Table 6

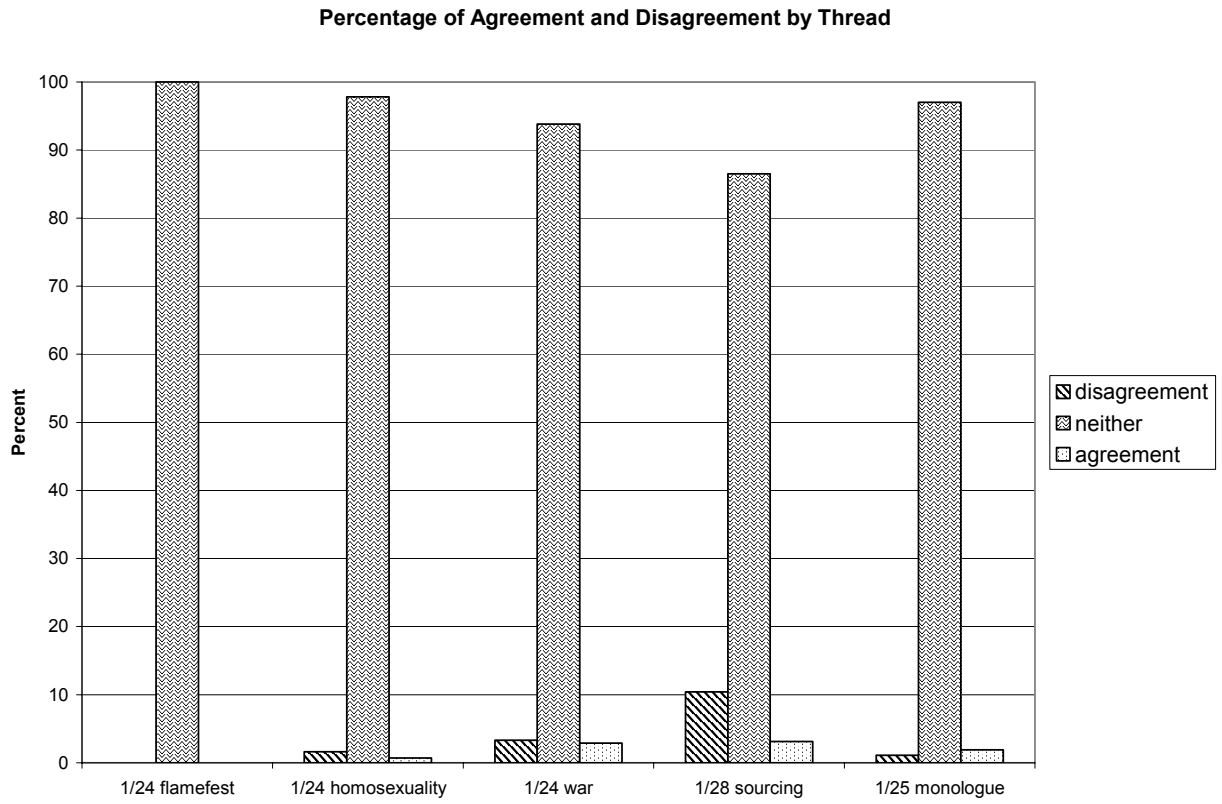


Table 7

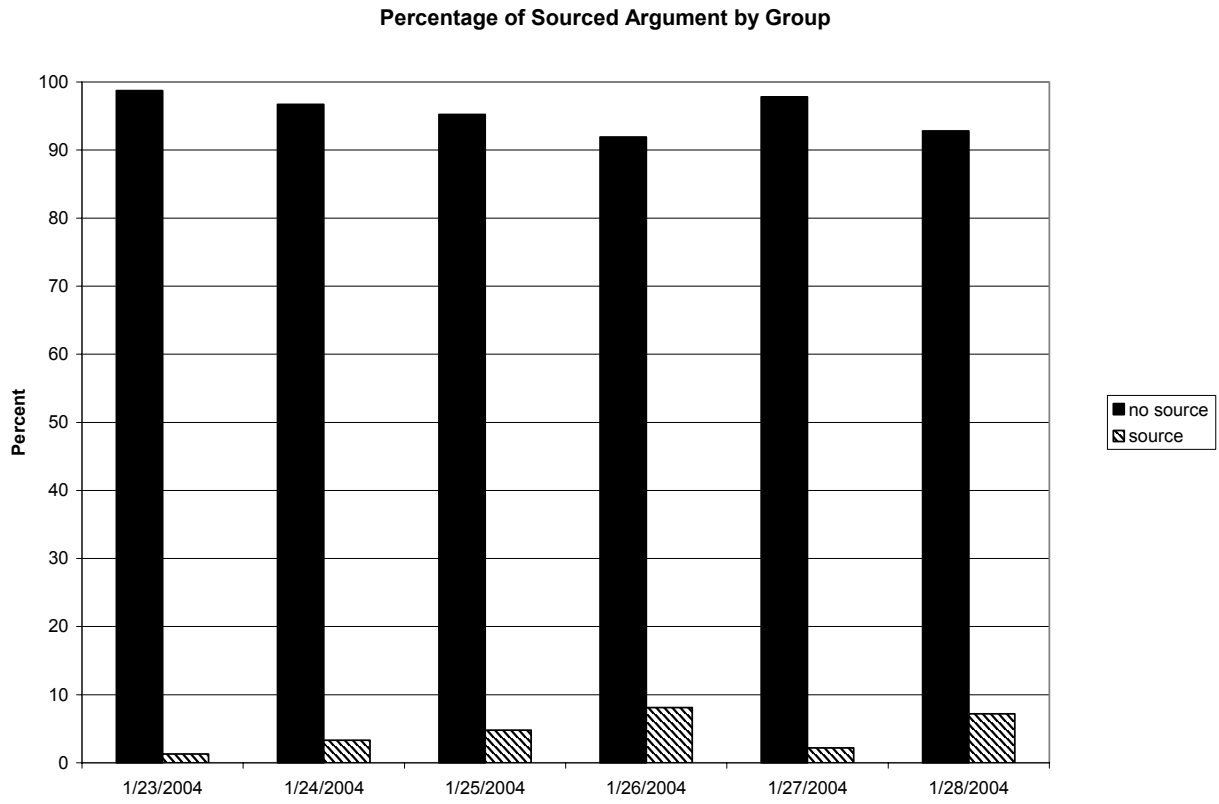


Table 8

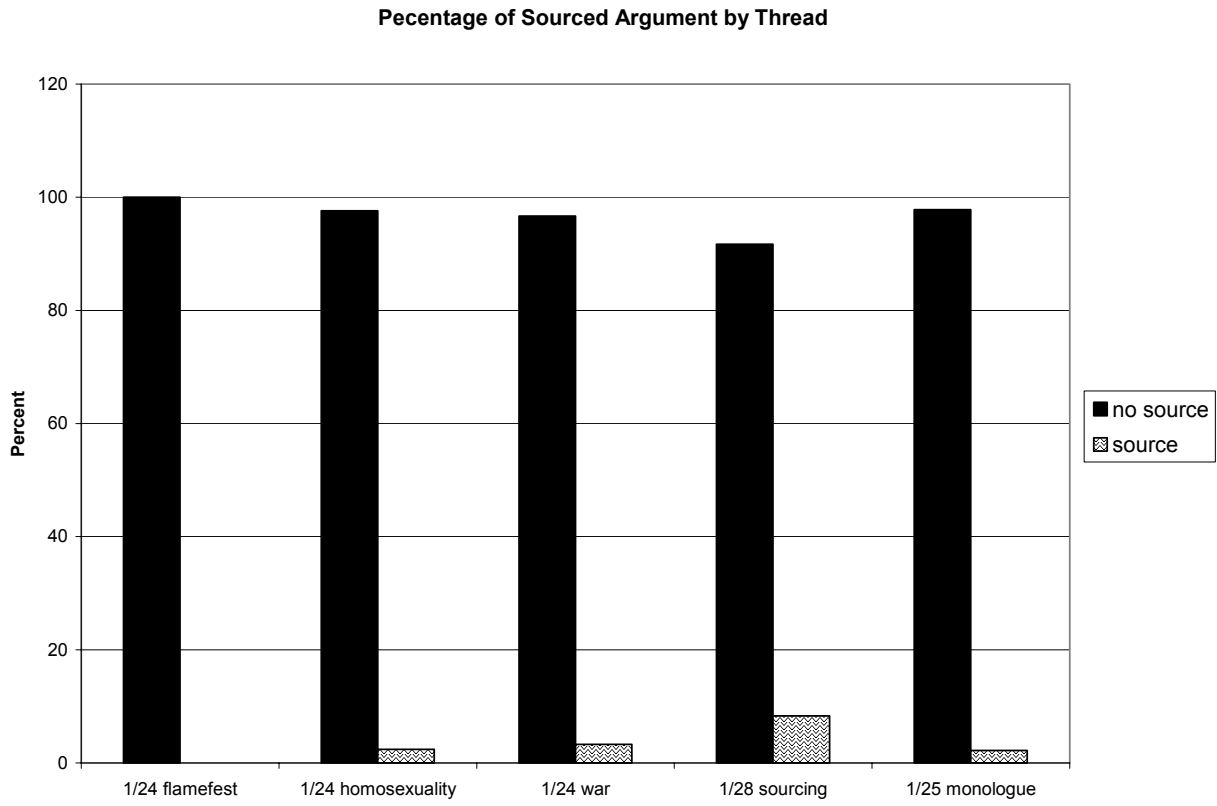


Table 9

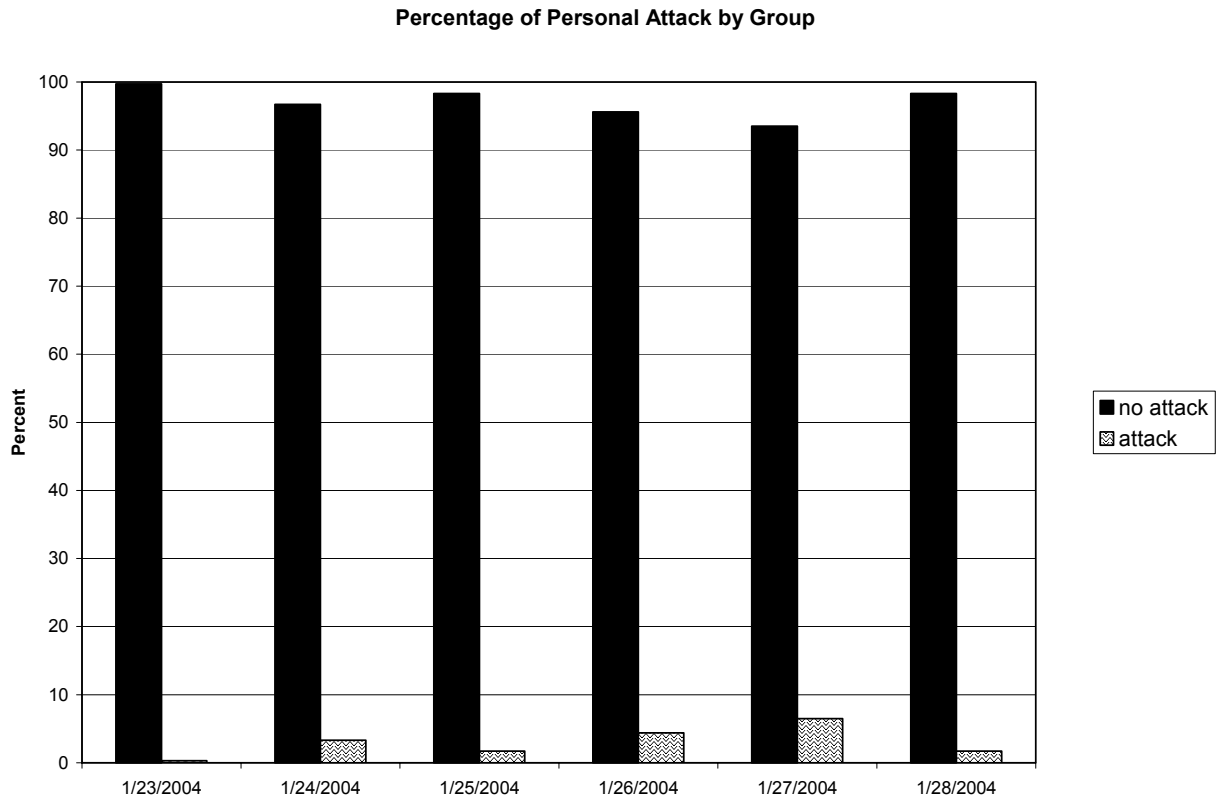


Table 10

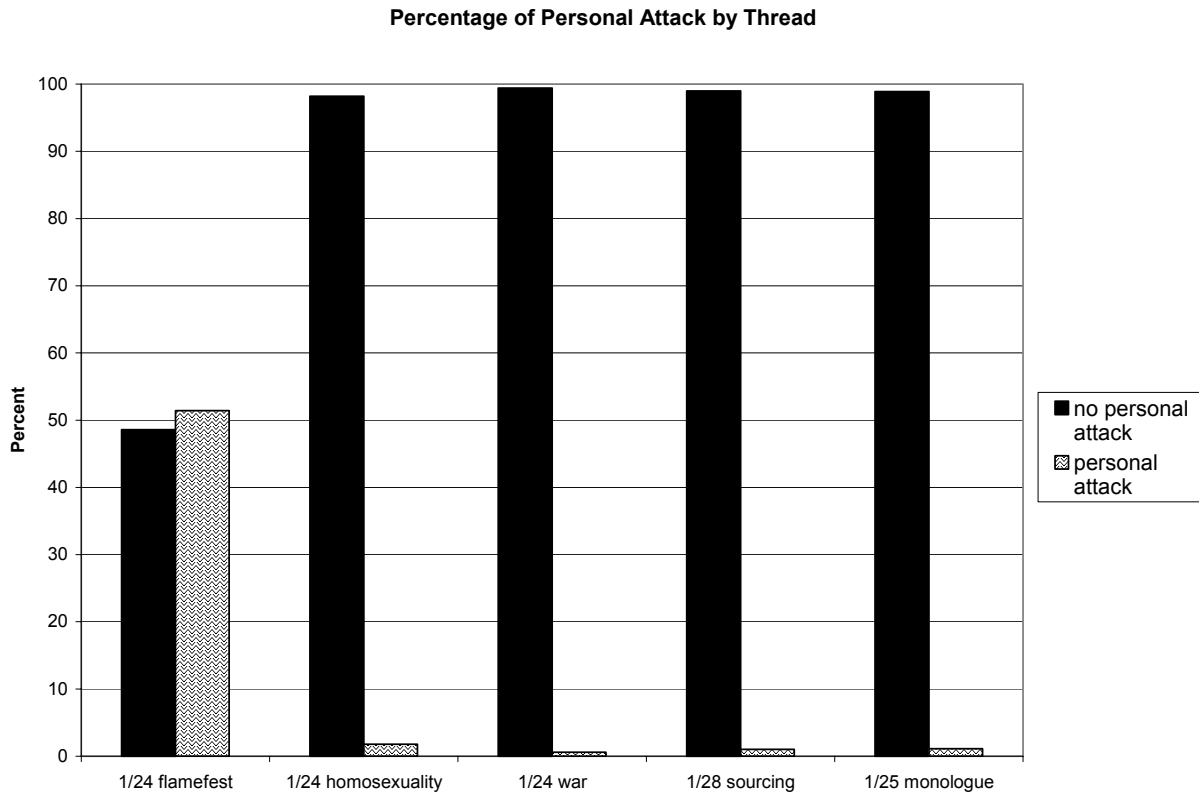


Table 11

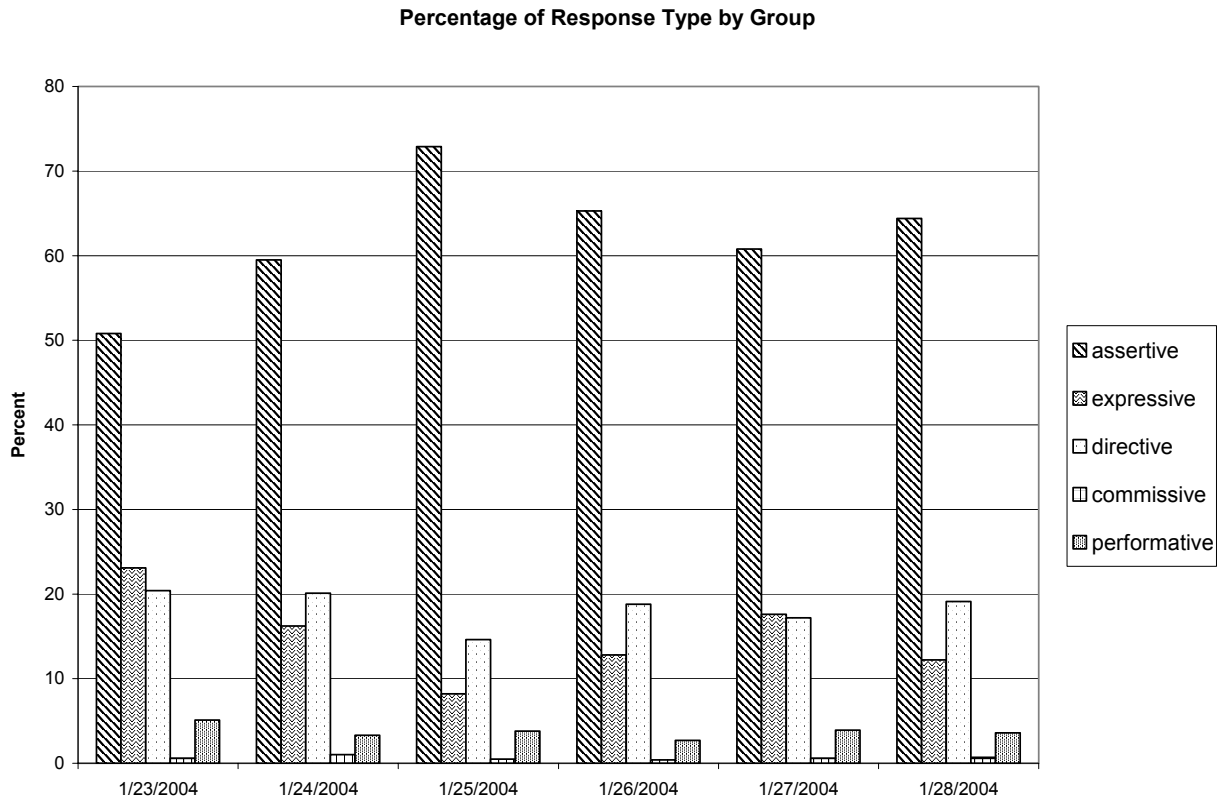
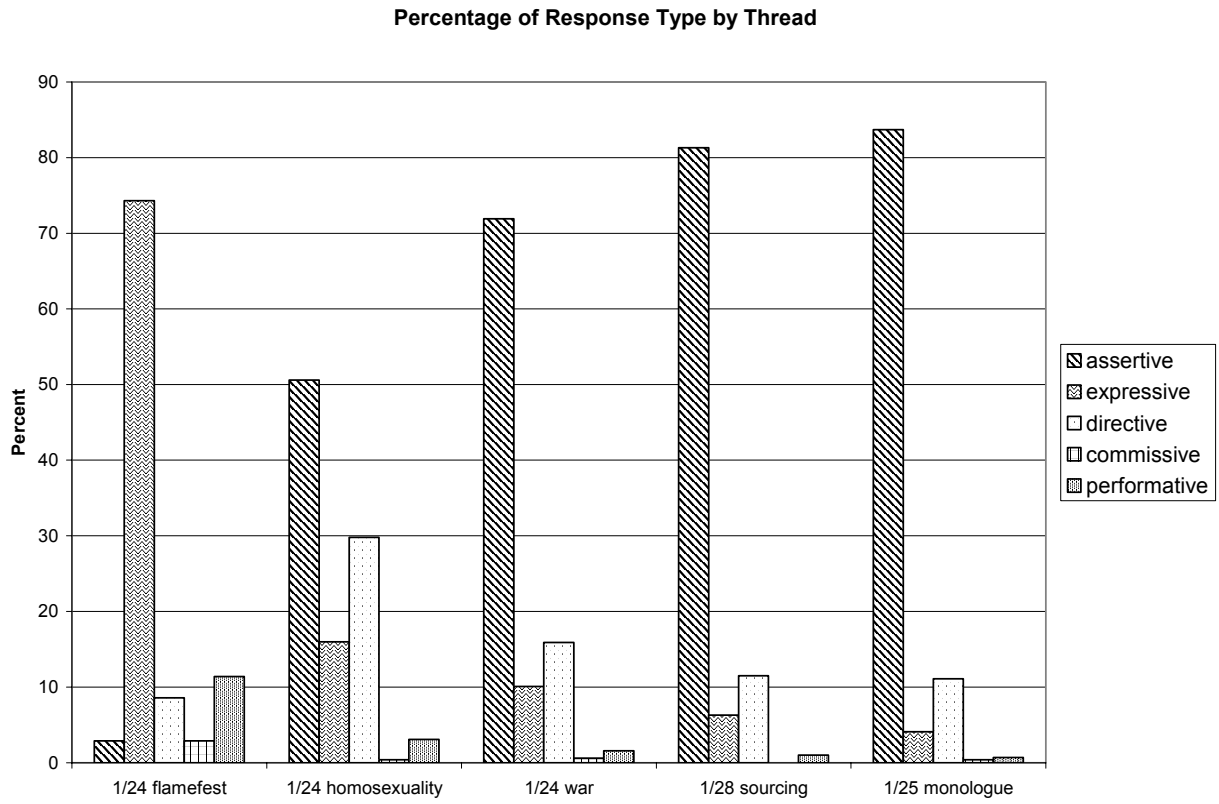


Table 12



References

- Burke, K. (1945). *A grammar of motives*. Berkeley, CA: University of California Press.
- Cappella, J. N., Price, V., & Nir, L. (2002). Argument repertoire as a reliable and valid measure of opinion quality: electronic dialogue during campaign 2000. *Political Communication*, *19*, 73-93.
- Cooren, F. (1999). *The organizing property of communication*. Philadelphia, PA: John Benjamins Publishing Company.
- Dahlberg, L. (2001). The internet and democratic discourse: Exploring the prospects of online deliberative forums extending the public sphere. *Information, Communication and Society*, *4*, 615-633.
- Davis, R. (1999). *The web of politics: The Internet's impact on the American political system*. New York: Oxford University Press.
- Doheny-Farina, S. (1996). *The wired neighborhood*. New Haven, CT: Yale University Press.
- Dutwin, D. (2002). *Can People Talk Politics? A Study of Deliberative Democracy*. Unpublished Dissertation, University of Pennsylvania, Philadelphia.
- Fishkin, J. S. (1991). *Democracy and deliberation: New directions for democratic reform*. New Haven, CT: Yale University Press.
- Fishkin, J. S. (1995). *The voice of the people: Public opinion and democracy*. New Haven, CT: Yale University Press.
- Gamson, W. A. (1992). *Talking politics*. Cambridge: Cambridge University Press.
- Gastil, J., & Dillard, J. P. (1999). Increasing political sophistication through public deliberation. *Political Communication*, *16*, 3-23.

- Graham, T., & Witschge, T. (2003). In search of online deliberation: Towards a new method for examining the quality of online discussions. *Communications*, 28, 173-204.
- Habermas, J. (1984). *The theory of communicative action*. Boston, MA: Beacon Press.
- Hill, K. A., & Hughes, J. E. (1998). *Cyberpolitics: Citizen activism in the age of the Internet*. Lanham, MD: Rowman & Littlefield Publishers, Co.
- Lemus, D. R., Seibold, D. R., Flanagin, A. J., & Metzger, M. J. (2004). Argument and decision making in computer-mediated groups. *Journal of Communication*, 54(2), 302-320.
- Luskin, R. C., Fishkin, J. S., & Iyengar, S. (2004). *Considered Opinions on U.S. Foreign Policy: Face-to-Face versus Online Deliberative Polling*. Paper presented at the International Communication Association, New Orleans, LA.
- Neuendorf, K. A. (2002). *The content analysis guidebook*. Thousand Oaks, CA: Sage.
- Price, V., Cappella, J. N., & Nir, L. (2002). Does disagreement contribute to more deliberative opinion? *Political Communication*, 19, 95-112.
- Ryfe, D. M. (forthcoming, 2005). Does deliberative democracy work? In N. Polsby (Ed.), *Annual review of political science* (8th ed.). Palo Alto, CA: Annual Reviews, Inc.
- Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Searle, J. R. (1979). *Expression and meaning: Studies in the theory of speech acts*. Cambridge: Cambridge University Press.
- Sunstein, C. R. (2001). *republic.com*. Princeton, NJ: Princeton University Press.
- Sunstein, C. R. (2003). The law of group polarization. In J. S. Fishkin & P. Laslett (Eds.), *Debating deliberative democracy* (pp. 80-101). Oxford: Blackwell Publishing.

- Twitchell, D. P., Adkins, M., Nunamaker, J. F., Jr., & Burgoon, J. (2004). *Using speech act theory to model conversations for automated classification and retrieval*. Paper presented at the Proceedings of the 9th International Working Conference on the Language-Action Perspective on Communication Modelling, New Brunswick, NJ.
- Vanderveken, D. (1990). *Meaning and speech acts* (Vol. 1). Cambridge: Cambridge University Press.
- White, C. S. (1997). Citizen participation and the Internet: Prospects for civic deliberation in the information age. *The Social Studies*, 88, 23-28.

Endnotes

¹ Observations of the forums suggest that during the day and the late evening hours the chat rooms are occupied with Europeans and Australians as well as Americans. This data was originally recorded to analyze political talk about the New Hampshire primary; however, there was very little sustained talk about them, so the recordings were instead used for this pilot project.

² Due to time constraints, those calculations have not been provided for this paper.

³ it would be worthwhile to calculate the number of participants who are active in the discussion and see if the interaction includes several people or only a couple. Due to time limitations, however, that calculation has not been conducted for this paper.

⁴ It is not necessarily the case that questions keep the conversation on topic. It is possible that questions function to shift the conversation to other topics, as well.