Argument Quality and Deliberation on Facebook: An Exploratory Study

Jessica Welch

In the past, scholars have argued that deliberation is important for a variety of reasons, including aiding in the development of political sophistication and decreasing attitudinal uncertainty (Gastil and Dillard 1999, 4-5). Engaging in deliberation allows individuals to become more informed about topics, which helps them to make informed decisions. This paper argues that if high argument quality is one requisite for deliberation, then deliberation is not occurring on Facebook due to its users’ low-quality arguments. In order to test this claim, 71 Facebook conversations related to President Trump’s immigration policies were analyzed. Results showed that users’ arguments were generally low quality and confirmed previous research that states that effective deliberation does not occur on social media sites. Potential consequences of these results and avenues for future research are subsequently discussed within this paper.

Jessica Welch is a Ph.D. candidate at Purdue University. She studies communication technology, focusing specifically on social media interactions.
Literature Review

Deliberation is defined as a communicative process in which groups engage in rigorous analysis of an issue and participants are attentive and carefully weigh the reasons for and against arguments (Black, Burkhalter, Gastil, and Stromer-Galley 2013, 3; Halpern and Gibbs 2013, 1160). Deliberation has many benefits at both the individual and societal level. On an individual level, deliberation can increase political sophistication and decrease attitudinal uncertainty (Gastil and Dillard 1999, 4-5). At the societal level, it results in the formation of better ideas and subsequently, better decision making by citizens (Cappella, Price, and Nir 2002, 75-77).

Unfortunately, recent research (Stroud, Scacco, Muddiman, and Curry 2015, 188) has found that, particularly on social media, conversations are not living up to the ideals of deliberation. These ideals include civility, relevant comments, asking genuine questions, and providing evidence to support your position (Stroud et. al. 2015, 190). Individuals engaging in ideal deliberation are also attentive and carefully weigh the reasons for and against arguments (Black et. al. 2013, 3). One explanation for the lack of deliberation on social media platforms is the low quality of arguments. Therefore, this study analyzes conversations on two political pages in order to determine the quality of arguments in Facebook comments, whether deliberation is occurring on social media, and the extent to which argument quality impacts deliberation.

Quality Arguments as a Requisite for Deliberation

As previously stated, deliberation is a communicative process in which groups of people engage in the rigorous analysis of an issue and attentive participants carefully weigh the reasons for and against arguments (Black et. al. 2013, 3). The process should include building an information base, prioritizing key values, identifying and weighing solutions, and coming to the best possible conclusion (Black et. al. 2013, 4). According to Cohen (2003, 347), another element of ideal deliberation is that all arguments must be supported with evidence. He writes that individuals should commit to solving problems through reasoning and that deliberation only occurs if the outcome results from free and reasoned arguments. Arguments are “reasoned” if the individual can provide a logical explanation for why they support or criticize the idea (Cohen 2003, 349). The best explanations are objectively verifiable, meaning that they have truth values that can be proved or disproved (Park and Cardie 2014, 31).

The problem is that conversations taking place on social media sites do not always live up to these ideals (Stroud, et al. 2015, 188). Because quality arguments are a necessary element of deliberation, the use of low-quality arguments on social media may be one reason that deliberation is not taking place. An argument is defined as a sequence of relevant and true premises that support a conclusion (Walton 1990, 400). Therefore, a low-quality argument would be one that includes irrelevant, false, or misleading premises, or statements that do not clearly support the conclusion.

Importance of Deliberation

Democratic theory assumes that voters will learn about their leaders’ policy positions before electing them, but previous research shows this may not be the case (Cappella et. al. 2002, 74). One cause of this issue may be a lack of effective deliberation about
politics. Research shows that engagement in deliberation benefits society in a variety of ways. One benefit of deliberation is that it increases political sophistication (Gastil and Dillard 1999, 4-5). According to Fiske and Taylor (1984) political sophistication can be defined as a “cognitive structure that represents one’s general knowledge about a given concept” (13). Deliberation improves this cognitive structure by increasing individuals’ schematic integration and differentiation, while decreasing their attitudinal uncertainty (Gastil and Dillard 1999, 4-5). Individuals’ beliefs demonstrate integration and differentiation if they are ideologically consistent; low attitudinal uncertainty is apparent if they are able to confidently and clearly state their opinion on an issue. Another benefit of deliberation is that it encourages people to reflect on issues and engage in critical thinking—which then results in the formation of well-reasoned opinions. (Cappella et. al. 2002, 74).

Measuring Argument Quality

In order to demonstrate that the comments posted on Facebook are low-quality, a conceptualization of argument quality must be developed. Many past attempts have made significant contributions to the field but—due to a lack of clarity and consistency (Boller et. al. 1990, 321)—researchers have yet to reach a general consensus on the best way to measure and define argument quality (O’Keefe and Jackson 1995, 88). According to O’Keefe and Jackson, there are three primary approaches to the operationalization of argument quality: pre-test procedure, argument quality ratings, and unsystematic message variations (1995, 88). In the pre-test procedure, study participants rate potential arguments for persuasiveness. The issue with this method is that participants are rating persuasiveness rather than quality. Fallacious and irrelevant statements may be persuasive in some instances, but still do not represent high-quality arguments. In the argument quality rating approach, participants rate the overall quality (rather than just the persuasiveness) of arguments. A limitation of this approach is that ratings are based on participants’ perceptions of what makes a good argument. People likely interpret argument quality in various ways, so ratings would be inconsistent across participants. Finally, during unsystematic message variation, the researcher manipulates messages based on characteristics that they believe influence the quality of arguments to determine which characteristics participants perceive as more effective. Again, this system is based on the researchers’ and participants’ perceptions of argument quality and will vary based on who is rating the message. Therefore, a more objective measure of argument quality must be developed.

In order to create a new conceptualization of argument quality, this study adopts and combines two conceptualizations developed by past research. Primarily, this new operationalization will be based on Boller et. al.’s (1990, 322-23) four crucial elements of an argument. These four elements include: claim assertions, evidence, authority, and probability. The first element—claim assertion—refers to the extent that an individual can effectively and clearly state their argument (322). For example, an individual who comments, “I agree with the travel ban” or “I think the travel ban is unconstitutional,” is using good claim assertion because you can clearly tell what their position is. The second element of argument quality—evidence—refers to the reasons and support that the individual provides to back up their argument. Based on a study conducted by Cappella et. al. (2002, 77), this means that the supporting
reasons must be relevant. For example, the comment “I disagree with the travel ban because it would cost the U.S. money” uses the potential for economic loss as the evidence for their argument.

The third element of argument quality—authority—refers to “warrants and backing,” which is when the speaker connects the evidence to the claim and demonstrates how they are related (Boller et. al. 1990, 323). For example, a comment that uses authority could read, “I disagree with the travel ban because it would cost the U.S. money. If we limit the people that can visit the United States, we will lose money from tourism.” In this statement, the potential for losing money is backed by the likelihood of decreased tourism under a travel ban.

Finally, the fourth element—probability—represents qualifiers and rebuttals. It refers both to the extent that individuals are willing to admit their claims are not absolute and to their ability to refute opposing arguments. For example, an individual could qualify their position by saying “I disagree with the travel ban unless we have documented proof that it will make our country safer.” Someone may also write, “You say that the travel ban would cost our country money in tourism, but very few people from the countries on the banned list come to the U.S. for vacation.” In both instances, the comments demonstrate probability because they either qualify or refute a claim.

The last study incorporated into the current conceptualization of argument quality is Hornikx and Hahn’s research on the frequent use of fallacious arguments in discourse. They claim that fallacies—specifically argument ad hominem—occur often and are generally not noticed (2012, 233). Fallacies are statements that violate a procedural norm of a rational argument (236). For example, in the ad hominem fallacy, an individual violates the rules of argumentation by dismantling their opponent rather than their opponent’s argument. Specifically, they use personal attacks to make their opponent seem less credible, rather than finding fault in the opposing argument itself. Walton claims that the use of fallacies indicates an erroneous argument (1990, 399). Therefore, in the present study, the presence of fallacies in any Facebook comment will negatively affect the rating of that argument’s quality.

**Measuring Deliberation**

Previous research indicates that, in order for a discussion to be considered deliberation, all comments must be civil, relevant, and not misleading (Coe et. al. 2014, 658-59; Stroud et al. 2015, 190). Unfortunately, several scholars have discovered that conversations occurring on social media rarely fit the requirements of ideal deliberation (Coe et. al. 2014, 658-59; Diakopoulos and Naaman 2011, 4-9). Therefore, effective deliberation in this study will be measured based on whether comments in each conversation are civil, relevant, and not misleading. Civil comments are those that do not contain profanity, threats, or name-calling. Comments are relevant if they contain only information about the post or introduce additional information that is pertinent to the post. For example, if an individual comments on a post about President Trump’s travel ban and claims that the Obama administration created the list of banned countries, that is a relevant comment. However, a comment that discusses the improvements that President Trump has made to the economy is irrelevant. Finally, a misleading comment includes information that is false or tries to pass an opinion off as fact.
Based on the previously reviewed literature, the following hypotheses and research questions are posed (see Model 1):

RQ1: What is the overall quality of the arguments used in these Facebook conversations?

RQ2: How often does deliberation occur in these Facebook conversations?

H1: Conversations including higher-quality arguments are more likely to be deliberative than conversations including low-quality arguments.

Model 1. Visualization of Argument Quality Measurements and Hypothesis 1

Methods

Conversations were collected from both the official Democratic and Republican party Facebook pages to avoid bias based on party affiliation. The dataset includes four posts (two from each page) related to President Trump’s immigration policy with comments and replies to those comments. Only posts regarding President Trump’s travel ban (often referred to as the “Muslim ban”) were collected to prevent variance that may occur between topics. Posts and comments were collected in spring 2017. Data collection began in January, less than two weeks after President Trump signed the executive order banning individuals from seven predominantly Muslim countries from visiting the U.S. for 90 days. This topic was chosen because of its timeliness and controversy. As previously stated, I wanted to limit conversations to one topic and a timely topic guaranteed sufficient posts and comments for coding. Second, controversial issues typically encourage individuals who disagree to discuss their views on Facebook in conversations that range from deliberation to virtual shouting matches. I wanted conversations wherein people discussed their disagreements in a variety of ways. In total, 71 conversations, including 328 individual comments, were collected and coded for argument quality and deliberation. Comments were evenly distributed across political parties.

Individual comments were coded based on argument quality using Boller et. al.’s (1990, 322-23) four crucial elements of an argument (claim assertion, evidence, authority, and probability) and Hornikx and Hahn’s (2012, 233) research on the frequent use of fallacies in arguments. Specifically, each comment was coded for whether the individual: 1) clearly stated their opinion, 2) provided evidence, 3) explained how that evidence support their opinion, 4) qualified their opinion or refuted another opinion, and 5) avoided fallacious reasoning (specifically, avoided use of argument ad hominem). Argument ad hominem was the type of fallacy I coded for because Hornikx and Hahn (2012, 235) found that it is the most common fallacy in social media conversations. Each variable was coded dichotomously (0=not present; 1=present). In this way, each comment received a score ranging from zero to five with a higher number representing a better argument.

For example, the comments in Images 1 and 2 both received a score of zero for argument quality. Neither comment includes a clear claim assertion, evidence, authority, or probability. In other words, the authors of
both comments fail to clearly and effectively articulate their position on the issue, they do not provide evidence or indicate how that evidence supports their opinion, and they do not include a qualifier or rebuttal. Furthermore, both comments include an ad hominem fallacy (personal attack). It can be implied that the writer of comment one supports the travel ban, while the individual who wrote the second comment is against it, but points for claim assertion were only awarded when the opinion was clearly stated.

Image 1. Low-Quality Argument

Image 2. Low-Quality Argument

Although far from perfect, Image 3 represents a better argument. The comment gets one point for claim assertion because you can clearly identify the individual's opinion (President Trump is on solid legal ground). The author also provides evidence in the form of an excerpt from the Immigration and Nationality Act (one point). Finally, the author avoids the use of fallacies (one point).

Image 3. Better Quality Argument

In order to determine the level of deliberation that occurred, each comment in a conversation was coded based on whether it was: 1) irrelevant, 2) uncivil, and 3) misleading. These variables were coded dichotomously (0=not present; 1=present). Comments that received zeros for items 1-3 were considered most deliberative. Conversations consisted of a comment on a post and the replies to that comment. It was not necessary for any individual to comment more than once for it to be considered a conversation. The more comments in a conversation that were coded as irrelevant, uncivil, or misleading, the less deliberative it was.

For example, Image 4 is an example of a conversation that was coded as not deliberative. The first comment was irrelevant because it mentioned homelessness, veterans, foster children, and abortion. It was also misleading by implying that preventing Muslims from coming to America would stop terrorism. The second and third comments were uncivil because they included personal attacks.
Image 4. Low Deliberation Conversation

Image 5 is a conversation that is closer to deliberation but was still coded as not deliberative due to the dichotomous nature of the coding (each conversation is either coded as 1, indicating that deliberation occurred, or coded as 0, indicating that deliberation did not occur). It includes some comments regarding the nature of the media and the travel ban that could be considered misleading, but the conversation remained civil and all comments were relevant.

Image 5. Moderate Deliberation

Results

To answer RQ1: “What is the overall quality of the arguments used?” frequencies were calculated to determine how many conversations included comments with each of the four crucial elements of argument quality (Boller et. al. 1990, 322-23). Of the 71 conversations analyzed, 58 (77.3%) included claim assertions (the individual clearly stated their argument), 32 (42.7%) included evidence (the argument was backed up with reasoning and evidence), 3 (4%) included authority (the individual clearly demonstrated how the evidence supports their argument) and only 1 (1.3%) included probability (the comment included a qualifier or rebuttal). Regarding fallacy use, 42 conversations (56%) included at least one fallacy. Therefore, in response to
RQ1, results indicate that the overall quality of arguments in this dataset was low. In fact, only one conversation (1.3% of the sample) met all the requirements for a high-quality argument and included a claim assertion, evidence, authority, probability, and fallacy-free) and was coded at the comment level. For example, arguments that included several pieces of evidence were considered higher quality than an argument with only one piece of evidence. Similarly, arguments with multiple fallacies were considered lower quality than arguments including only one fallacy. Results indicate that in every instance in which deliberation occurred, arguments were fallacy-free and included at least a claim assertion, with the majority also including evidence. On the other hand, conversations in which deliberation did not take place generally included multiple fallacies and were missing the four crucial elements of argument quality (Boller et. al. 1990, 322-23). In some cases, the “evidence” element was present, but in others the element was missing. In other words, the individual provided support for an argument that they did not clearly articulate. Overall, the conversations that were considered deliberation included higher-quality arguments than those that were not. Thus, Hypothesis 1 was supported.

One interesting result that was not predicted by the hypotheses or research questions was that the use of inflammatory language in the original post did not impact the argument quality or deliberation of conversations. For example, comments on the Democrat’s post calling the travel ban “morally bankrupt” were no less civil or relevant than comments on their less provocative post. Additionally, the Republican post that mentioned 9/11 as an excuse for the travel ban did not elicit more incivility or lower quality arguments. In other words, comments were consistently low quality and conversations lacked deliberation regardless of the original post.
Discussion and Conclusions

This exploratory study supports the claim that higher-quality arguments lead to more deliberative conversations. Results also confirm previous research on deliberation which found that it rarely occurs on social media sites. This analysis indicated that—as Hornikx and Hahn (2012, 235) found—argument ad hominem (name calling) is all too common in online discussions. This name calling is one example of the lack of civility on social media, which leads to a lack of deliberation. Deliberation, defined as a communicative process in which groups engage in rigorous analysis of an issue (Black et. al. 2013, 3), helps individuals to become more informed about topics by learning from others’ opinions and experiences. The fact that individuals have so many opportunities to exchange information and engage in civil dialogue with others, and yet rarely do so, may be a symptom of a larger societal problem.

One theory that may shed light on individuals’ lack of ability or desire to engage in productive deliberation is Motivated Reasoning. Motivated Reasoning is a goal-directed strategy for cognitive processing in which individuals seek out information that confirms their prior views, consider evidence consistent with their opinions as stronger, and spend time arguing against evidence inconsistent with their opinions (Druckman 2012, 200; Nir 2011, 505-6). According to this theory, when individuals are faced with new information, their analysis of it is biased based on their previous beliefs. This is an obstacle to deliberation because—when faced with information that contradicts their position—individuals may ignore or discount it rather than updating their views (Nir 2011, 505-6). Motivated Reasoning is particularly common with highly partisan topics (like the travel ban) when individuals feel pressure to agree with the dominant position of their political group. In these cases, party allegiance may be stronger than an individual’s opinion on the topic, leading them to maintain their position even when faced with contradictory evidence (Gaines et. al. 2007, 963). In this way, discussion about a political issue may lead to more polarization between groups rather than compromise (Hart and Nisbet 2011, 702-5). Therefore, the low-quality arguments and lack of deliberation found in this research may be partially explained by the political topics studied. Future research should examine argument quality and deliberation surrounding a variety of topics to determine how results might differ.

In addition to studying comments on varying topics, there are some other limitations that future research could address. Future research should examine dialogue on both a variety of topics and on a variety of Facebook pages. The Facebook users who visit the pages examined in this study likely hold stronger political views and may therefore be more close-minded when it comes to discussing controversial issues. Conversations taking place on different pages may be less extreme, include higher-quality arguments, and be more comparable to ideal deliberation. Another issue that previous research has encountered, and that this study also experienced, is the operationalization of argument quality. Research still lacks a general characterization of argument quality and agreement on what elements are necessary to high-quality arguments (O’Keefe and Jackson 1995, 88). This study attempts to improve previous measures by combining two of the most parsimonious characterizations (Boller et. al. 1990, 322-23; Hornikx and Hahn 2012, 232-38), but the problem—lack of consensus on argument quality measures—remains. In this study, the argument quality
of each comment was determined by a single coder, so reliability could not be calculated. To validate the argument quality measurements I proposed, a future study could include a sample of the low, medium, and high-quality arguments from this study and ask participants to rate their quality without introducing them to the elements identified by Boller et. al. (1990, 322-23). Future research should also duplicate this study with several coders, rather than a single coder, in order to achieve reliability and produce more rigorous results. Despite these limitations, this study provides important contributions in the areas of argument quality, deliberation, and online engagement.
Argument Quality and Deliberation on Facebook

Bibliography


