1. Abstract

This paper develops a coding scheme for measuring deliberation in online, asynchronous communication of social news websites, in the context of the 2008 US elections. Certain qualities of deliberation (the heterogeneity of expressed views, the argumentativeness of the discussion, and its reciprocity) are suggested to be measured through the analysis of the text of conversations themselves. The analysis of these qualities is illustrated on a sample discussion piece lifted from social news site Reddit.
2. Introduction

This paper describes a possible way to operationalize deliberation in an online environment. Relying on existing ways to measure the deliberative potential of conversations, it develops a coding scheme to analyze some aspects of online, asynchronous discussions on social news websites, in the context of the 2008 presidential election in the US. An important point this paper makes is that deliberation needs to be understood and analyzed in a specific context. Thus, it is more than likely that the coding scheme developed here needs to be modified, in case one intends to apply it in the analysis of conversations pertaining to different contexts. However, in the course of developing the coding scheme, I hope to raise important points about analyzing online deliberation in general.

In the first part of the paper, I briefly review some previous attempts aimed at defining and operationalizing deliberation. Relying on these attempts, I then develop a new coding scheme – which I demonstrate through the analysis of a sample piece of discussion lifted from a social news website.

This piece of research is one building block in my dissertation, in which social news sites are analyzed from an external perspective (as media organs), and from an internal perspective (as discursive, potentially deliberative, spaces). This paper in its current form is prepared exclusively for our research seminar.

3. Literature review

3.1. What is deliberation?

The precise definition of deliberation is often conspicuously absent from works that discuss ways of measuring it. It is often taken for granted that the meaning of the term is known and understood; that a more or less hazy and implicit definition – “some kind of serious, careful debate about something” – is enough. In contrast, I start from defining the concept. The Merriam-Webster (2009) online dictionary defines deliberation as “a discussion and consideration by a group of persons (as a jury or legislature) of the reasons for and against a measure.” This is a good starting point, but also one that is incomplete in some way. It identifies two of the three factors that generally feature the implicit conceptualizations of existing research.

First – at the heart of the concept of deliberation is the thought that several people contemplate together about something, trying to reach some kind of conclusion. The idea that the many, when put together, can reach better decisions than the individual dates back at least as far as Aristotle¹.

However, the communal decision-making of any given group about any given matter does not necessarily qualify as deliberation. The many, “when put together,” need to discuss, to exchange information in order to arrive at a decision. This is a second important feature of deliberation: that it does not consist of the mere aggregation of the many different views in a community. (In contrast, consider the idea of the “wisdom of crowds,” which suggests that simply calculating the average of the separate, individual and uninfluenced opinion of members of a group will produce a better decision than any member of the group would reach on his or her own. (Surowiecki 2005.))

What is missing from the above definition is that deliberation also implies that not any kind of discussion and information exchange does suffice. Deliberation is discussion that (and this is the third point) needs to meet particular criteria. Opinions differ over what exactly these criteria are, but

¹ “Now any member of the assembly, taken separately, is certainly inferior to the wise man. But the state is made up of many individuals. And as a feast to which all the guests contribute is better than the banquet furnished by a single man, so a multitude is a better judge of many things than any individual.” (Aristotle 2000: 136 – 137)
the main point is that deliberation, especially if it figures as a constitutive element of democracies, needs to be *good discussion*; a process which assures – as far as this is possible – that it will lead to a *good*, or at any rate, *better* decision.

So to recap, my concept of deliberation, distilled from previous research material, rests on the three elements of (1) inclusion of more than one person (2) discussing about a certain matter (3) in such a way that the discussion satisfies a certain set of criteria.

Existing attempts to evaluate the deliberative potential of discussions fit well into this framework: it is a generally accepted practice to establish what these aforementioned criteria are, and to look for signs in the discussion by which the extent to which these criteria are fulfilled can be measured.

But I would like to stress another important point: notably, that deliberation might mean different things in different contexts, and in relation to different problems (Habermas 2005: 387). Deliberation is hardly ever analyzed on its own, as some peculiar form of discussion. It is the idea of deliberative democracy that attaches weight to the concept: in models of deliberative democracy, deliberation is crucial because it *is the process that is expected to guarantee the legitimacy, justice or moral rightness of a political system.*

The perceived importance of deliberation often lead scholars to the creation of strict and formal deliberative criterions (e.g. Steenbergen et al. 2003, Dahl 1989, Cohen 1989 and others). These criterions consider deliberation as a monolithic concept, suggesting that there is only one way to correctly deliberate, that there is only a single “ideal process” which political decision making should follow. (The same assumption is often made by critics of deliberative models of democracy too, e.g. Sanders 1997.)

In accordance with Habermas (2005, 2006, 2009), I would like to take a different view, and stress that deliberation can refer to a number of different discursive practices, and that it has different roles in different contexts.

Naturally, this means that the same set of criteria may not be suitable for measuring deliberation in different circumstances. The “everyday deliberation” of ordinary citizens cannot be measured by the same standards as the formal, political and explicitly deliberative discussions of parliamentary committees or representatives.

Therefore, while relying on previous research in developing the coding scheme I intend to use for measuring deliberation in online discussions, I will consider the context of these discussions, and operationalize deliberation in accordance with this context.

### 3.2. What qualifies as deliberation?

Davy Janssen and Raphaël Kies (2005) provide a good starting point with their meta-analysis of “characteristics of the idealized public sphere.” Following the analytical framework developed by Dahlberg (2002), and analyzing the works of Dahlgren, Habermas, Schneider, Jensen, Jankowski, Van Os, Wilhelm, Graham and Steenbergen et al., they establish the following criteria for deliberation.

**Reciprocity:** only those discussions could qualify as deliberative in which parties take turns in talking and listening, and reply to each other, as opposed to leaving each other's points out of consideration and essentially delivering monologues. (In 3.1. I referred to the importance of deliberation being discussion – i.e. communication where the parties take in and respond to each other's claims.)

**Justification:** in deliberation, *reasons* that advocate a certain course of action are to be debated. If such reasons are not brought to the table, there may not emerge any kind of common ground on
which participants to the discussion can reach a consensus (or at any rate, an agreement of some sort). One might argue that the evaluation of the various reasons offered is the single most important idea of deliberation.

**Reflexivity:** whether or not deliberation “works” is indicated by whether it causes the participants' opinion to change (if not in direction, then in its extent or its backing). Conversants need to internalize the arguments put forward, evaluate them, and, if compelling arguments are offered, change their initial opinions accordingly. If this does not happen, i.e. if participants to a conversation do not give in to compelling arguments, then the conversation itself might be seen as useless from the point of view of deliberative decision making.

**Ideal role taking:** this characteristic refers to the ongoing and respectful listening that is judged, by some authors, necessary for deliberation. Ideal role taking means that participants try to consider all possible positions on an issue, not only the ones actually present at the discussion; that they treat each other with equal respect; and that they continue the discussion as long as it reaches some conclusion.

**Sincerity:** it is expected that participants to deliberation “must make a sincere effort to make known all relevant information, including their intentions, interests, needs, and desires” (Dahlberg 2002:3, as quoted by Janssen and Kies 2005: 329).

**Inclusion and discursive equality:** in deliberation, everyone potentially affected by its outcome should be able to participate (a feature to which we could refer as external openness). Further, it is also expected that all the participants have equal rights and opportunity to listen, and to be listened to (internal openness). Openness and inclusion is thought to be important, because it is through these qualities that we can expect all the relevant positions on an issue become represented in the discussion.

**Autonomy from state and economic power:** finally, deliberation should be free of the influences of political and economic power, representing instead the true, undistorted interests of the citizens.

Provided that we had ways to measure how a particular conversation stacks up against these ideals, we could decide – say Janssen and Kies – whether that particular conversation could qualify as deliberation; or at least evaluate to what extent it would approximate the ideal. The authors themselves draw attention to possibly valid criticism against the application of such a criterion in order to measure deliberation. Among other, mostly methodological, issues, they also make the point that the rigorous application of a fixed, pre-determined set of criteria to different kinds of discussion might be problematic, and might lead to invalid results (Janssen and Kies 2005: 331 – 332). With that in mind, I build upon their comprehensive overview in suggesting a slightly different conceptual approach.

4. Deliberation in a specific context

One cornerstone of my research is the Habermasian model of deliberative democracy in general, and in particular the following idea: deliberative political decision making in a society demands the cooperation of (a) the institutional political core, (b) the informal, wild flows of communication originating in the periphery that is civil society, and (c) the political and cultural public spheres that mediate between the former two.²

² “[The] sociological translation of the discourse theory of democracy implies that binding decisions, to be legitimate, must be steered by communication flows that start at the periphery and pass through the sluices of democratic and
Habermas refers to this as “two-track” deliberative politics. Formal political decision making takes place at the institutional centre – but it needs to rely on the informal and unorganized flows of communication of the periphery, which are crystallized, amplified or attenuated and pitted against each other in the public sphere (orchestrated by the media and influential public personalities). This arrangement suggests what Habermas also makes explicit: deliberation in the political core is different to the informal deliberation in the civil society and in the public sphere. Not only because of the different practical circumstances under which they take place, but also because of the fact that the informal, public deliberation of everyday actors is detached from the kind of political decision making that is the function of the political institutions. While the role of deliberation in the political core is to reach the best possible decision in administrative, legislative or judicial matters, the role of public, everyday deliberation is different: it is to make participants arrive at a considered opinion; to identify and draw attention to as many different perspectives as possible; and to influence how these different perspectives are narrated and debated in the public sphere.

On the other hand – and this is also a crucially important point of my research project – this model does not account for the case when the informal deliberation of everyday people is not detached from political decision making. This happens in the case of elections and referenda. One of the ideas behind elections and referenda is that they are a simpler solution than deliberation. There is no way in which the whole society could come together in some large (even if virtual) town hall and debate until reaching a consensus on whom to elect as president, or which way a question put to referendum should be answered. Instead, everyone just casts (or doesn't) his or her vote based on a large and varied number of reasons, and a decision will be made based on the aggregation of these individual votes.

But if we accept the idea of deliberative democracy, then in order for this decision to be legitimate, it needs to reflect the considered opinion of voters. The way in which each individual reaches his or her final opinion needs to be a “deliberative” way. This is – partly – where the responsibility of the public sphere lies. When political decision making is put in the hands of the citizens themselves (as opposed to some political decision making core), then the public sphere functions correctly only to the extent that it makes sure that voters deliberate over the matter they are supposed to decide on.

Admittedly, it might be utopian to expect the political public sphere to approximate this ideal. The point I'm trying to make here is twofold: first, that measuring public deliberation in instances of direct democratic decision making (such as elections) could provide an empirical “quality index” of the political public sphere; and second, that if this is so, then it is justified to set up a particular set of criteria (albeit a different one that is used in relation to formal political decision making), and measure the informal, public deliberation by it.

With this in mind, I suggest the analysis of public, informal deliberation in the context of the 2008 US presidential election. Considering this as an issue, from the point of view of the solution, the task is very simple. There are a number of candidates to choose from, two of which (candidates from the Republican and the Democratic parties) have a realistic chance of eventually becoming president of the USA (Ashbee 2004: 184 – 188, 227); and your task as a voter is to pick one of these candidates – the one that you think would do the best job as a president. There can be no compromise, “40-60” solutions; everyone has one vote, and this vote can only be cast on one of the candidates. The problem in abstraction is to find the best solution to the general question of “how to go on?”; but practically all it boils down to is the selection of one of the candidates.

---

constitutional procedures situated at the entrance to the parliamentary complex [...]” (Habermas 2009: 356; see also Habermas 2006 and 2009: 304 – 308).
In this situation, the role of deliberation is not more than making participants reach a rationally motivated, carefully thought out opinion, after having considered as many positions on the issue as possible. (Referring to Bohman (2007), a “position” here should mean not only the view on who to vote for, but it also should encompass the reasons for doing so. For instance, supposing that Hillary Clinton had been one of the presidential candidates, a feminist voter might support her for different reasons than a Mexican-American one – two different positions, with their respective sets of arguments for supporting the same candidate.)

In addition, I propose the analysis of public deliberation in the form of online, asynchronous communication flows: discussion threads attached to popular items published on social news websites. Consequently, my analytical tools will need to be adapted to this particular form of communication.

5. My criteria of deliberation

I arrived at the following scheme by trying to adapt existing operationalizations to the specific circumstances mentioned above. These are the qualities, then, which would describe public, online deliberation about the 2008 US presidential elections in an ideal form.

**Heterogeneity of views**

One of the most important ideas underpinning deliberation is that it introduces participants to a variety of views on the issue at hand. Successful deliberation requires that participants represents many different points of view (Mutz 2006, Sunstein 2007, Bohman 2007 and others). A problem often associated with online discussion forums is that they merely act as “echo chambers,” where participants only encounter views similar to their own, as a result of which not only do they miss potentially relevant contradicting arguments, but their views also tend to get stronger in extent (Sunstein 2007: 64 – 67), opening a way towards radicalization and extremism.

Yet existing analyses of online deliberation often fail to measure the existence of contradicting arguments, of opposing views (a notable exception can be found in Stromer-Galley 2005). What could explain this apparent contradiction?

One reason could be that existing schemes to measure deliberation try to focus on the discussion's procedural qualities. Procedural criteria are abstract, they do not depend on the actual contents of the discussion. The existence or lack of heterogeneous views in a discussion is, on the other hand, a measure related to the contents.

One might say that if the discussion is open, inclusive, and satisfies the condition of reciprocity, then it is reasonable to assume that a variety of conflicting, heterogeneous opinions is expressed in the course of discussion. In other words, we could infer information on the contents of discussion from certain procedural qualities.

But if this is true, I then suggest that we turn this chain of thought around, and instead of trying to establish the heterogeneity of views expressed from procedural qualities, let's try to deduce certain procedural qualities from the independently measured heterogeneity of views. Arguably, it is simpler and more commonsensical to affirm that the heterogeneity of discussions suggests an openness of the discussion, than to turn this statement around.

A first, very basic measure of heterogeneity of views in the specific context I'm looking at would be the measure of views supporting the Republican Party *versus* views supporting the Democratic Party (*versus* views supporting third parties).

**External openness**

The notion that deliberation should be open to everyone potentially affected by its subject matter is

---

3 Although it must be noted that most existing coding schemes do try to look for *rational arguments* in discussions, and this also involves the analysis of contents, in addition to the analysis of the procedure of discussion.
somewhat problematic when it comes to online discussions, for two reasons. First – if we take to understand the concept of “everyone potentially affected” in the original spirit of deliberative democracy – that it refers to all the citizens of a state, because the state is supposed to make decisions reflecting the best interest of its citizens –, then we are faced with a set of practical questions. How many citizens have actual access to the internet? How many of them are aware of the deliberative platforms that exist? Is there anything that could preclude certain groups of people from using these platforms?

The idea of mediated deliberation (cf. Xenos 2008) suggests that as long as all the relevant points of view are represented in the debate, it is not necessary for everyone to actually participate in the debate (this would be a practical impossibility (cf. Michelman 1997:154)). But the possibility of joining the debate should be open for everyone, in case it is noticed that a relevant point of view is missing from the debate.

Second – online deliberations can spawn across traditional country borders, and so openness could also mean that even such people will participate in them – in an undistinguishable manner – who will not be bound by the eventual decision reached in the actual act of voting. (This is true even if, in the case of presidential elections, citizens of certain other states have strong reason to argue that the outcome of the elections would influence their lives to much greater extent than the lives of Americans.)

However, I don't believe this would compromise the idea of deliberative democracy. For one thing, the truth, rationality and generalizability of an argument need not depend on the person that makes the argument, or on the citizenship of this person. Arguably, external perspectives could reveal relevant points of view that otherwise would be overlooked, as outsiders might also question such basic assumptions that are taken for granted by natives. And, importantly, while outsiders can and do participate in the debate, they are not going to participate in the actual decision making (voting), and so they will not compromise the idea of national sovereignty of political decisions.

In any case, ideal external openness would guarantee a discussion platform practically accessible at least to all American citizens.

Internal openness
A structural counterpart of external openness, this characteristic addresses the idea that every participant to the conversation should have equal opportunity to listen, and to be listened to. This is a standard feature also of existing operationalizations of deliberation.

Argumentativeness
It is trivial that deliberation cannot fulfill its main designated function, the evaluation of various reasons that justify various courses of action, in the absence of such reasons. In practice, this establishes a twofold condition that discussions should satisfy: first, that there be arguments offered by participants for their suggested claims, and second, that these arguments be valid – in the sense that there are good reasons for agreeing with them.

In the next chapter I describe in detail what I call an “argument,” and what would make an argument valid and relevant, in the specific context I'm looking at.

Reciprocity
In agreement with Janssen and Kies (2005), I also consider reciprocity as an important characteristic of deliberation. Reciprocity should mean that there is real conversation going on between the participants, i.e. that they listen to and reflect on each other's thoughts. This is especially important in an online setting where no practical circumstance obliges participants to listen to others.
Reflexivity
Ideally, deliberation should affect its participants in such a way that they adjust their opinion to comply with the new pieces of information and arguments that they were exposed to in the course of the conversation. If this doesn't happen, then deliberation remains, from the point of view of democracy, an entirely pointless activity.
But when speaking about the specific kind of deliberation I am studying, it must also be noted that it is not only the participants themselves that deliberation could affect. In an important contrast to deliberative polls (Fishkin 1991), this loose, online, informal deliberation is also public, openly accessible not only to the participants themselves, but also to those who prefer not to take part in the actual debate. (Just like discussion forums and blogs, social news sites are also full of “lurkers”: people who do not contribute to the discussion, only read what others have contributed.)

| To sum my conceptualization of deliberation up: deliberation – in the context of informal, online discussions related to the 2008 US presidential election on social news websites – is conversation that exhibits a **heterogeneity of views** about the topic, is both **externally** and **internally open**, is based on **valid arguments**, is **reciprocal**, and in which participants **reflect** on the proposed arguments and modify their own points of view on the subject accordingly. |

(A note on authenticity and autonomy. As you might have noticed, the categories of “sincerity” and “autonomy from state and economic power” are missing from my conceptual description of deliberation. As I will explain in the next chapter, my concept of valid arguments rests on Habermas' idea of communicative rationality, which in itself incorporates the idea of authentic, undistorted, autonomous messages. Therefore I believe that the establishing separate analytical categories for the measurement of these qualities would further complicate the analysis of deliberation, without bringing substantial new information to the table.)

6. Measurement method

Some of the proposed characteristics could be measured through the analysis of the contents of written and archived text of the debates themselves. These characteristics are the **heterogeneity of expressed views**, and their **justification** (argumentativeness).
Some, on the other hand – namely: the qualities of **openness** and **reflexivity** – should be measured not through inferences from the text of the debate, but by turning directly to the participants through a survey or interview.
Finally, the quality of **reciprocity** could also be inferred from the archived texts of the discussions, but not directly from their contents, rather, from their structural organization.

There are no established standards of deliberation: no certified and generally accepted claim can be made along the lines of “if a particular text contains 10 or more arguments, then it should be considered as deliberative.” What could be measured, on the other hand, is the extent to which actual discussions differ from an ideal situation. E.g. it could be proposed that in an ideal scenario, every speech act in a deliberation would contain at least a single complete argument, and it could then be measured what percentage of speech acts **actually** contain at least a single complete argument. This should provide a good basis for evaluating the deliberative potential of social news sites – but the development and calibration of an independent scale of deliberation should be the objective of subsequent research.

6.1. Measuring reciprocity

Conversations on social news sites are organized in the following manner.
There is a “comment space” attached to every single published article. In this comment space, various threads of discussion can develop: users can comment on the article itself, or, opening a branch or sub-thread, they can also reply to other comments (thus a thread refers to all the replies that users post either to the conversation starting article, or to any other subsequent comments).

It is common practice (Janssen & Kies 2005: 326, Stromer-Galley 2005: 11) to measure the reciprocal quality of discussions through counting the number of replies individual posts receive. This is a suitable, if not perfect, method for the analysis of discussion boards and online chat, where discussion threads cannot split up to form additional sub-threads. On a discussion board, a message is either a topic starter, or a reply to a topic starter message, and there is no structural difference between various replies. (Figure 1.)

Thus I propose the use of a scheme developed by Kaltenbrunner, Gonzalez-Bailon and Banchs (2009). Starting from social network analysis, this scheme treats branching discussion threads in online communication as networks, where the nodes are individual comment entries, and the links between them are established by replies. The scheme then measures two structural features of these networks: the average depth of the various branches of the threads, and their width, meaning the number of different sub-threads within a single conversation.
In the interpretation of Kaltenbrunner et al., depth could be referred to as the “intensity” of conversation, while width informs about the “interest” attracted by a discussion. Based on the ratio of a discussion network's depth to its width, the authors establish four different types of network: that of high depth and high width (Type 1), that of high depth but low width (Type 2), that of low depth but high width (Type 3), and that of low depth and low width (Type 4).

Kaltenbrunner et al. were less interested in the objective, normative qualities of discussions, and in deciding what is “high” and “low” in terms of depth and width, they took average measures of several analysed discussion networks. This method can hardly be applied in my case, since I am looking for some independent standard to judge the quality of discussions. Establishing a theoretical ideal situation could be a further task in my research project. For this paper, I merely attempt to record the ratio of the width and depth of conversations, and try to interpret these intuitively.

6.2. Unit of content analysis

Threads in the discussion can identify overarching themes, or topics of the discussion, but naturally they can further be broken down to smaller units, namely the individual posts or comment entries of which they consist. One particular user might, naturally, contribute to a discussion with one or many posts, and in this latter case, these posts are considered in separation.

Although one might argue that there is a sense in establishing a thread as the unit of analysis, or establishing individual narratives (all posts taken together from a single user) as the unit of analysis, I have decided not to do so, for the following reasons.

First, the individual comment entry is the smallest self-contained structural component of the discussion. In the course of a discussion, a person might contribute with comments that contradict each other – a change in opinion might reflect the influence of another comment, for instance –; while this is not expected to happen within a single post.

Second, social news sites themselves treat the post as the unit of discussion. Communication on social news sites is asynchronous: there is no option of real-time exchange of chat messages; and this means that the unit of discussion cannot be, structurally, anything else then a single comment entry. In addition, sites often also enable users to rate each other's posts, and the handling and branching of discussion threads also rests on treating the separate posts as the basic unit of conversation.

Third, the sheer number of posts can hold information that might be lost when choosing a larger unit of analysis. In a trivial example: a discussion in which two distinct lines of thought or

![Figure 2. Structure of discussion networks. (Kaltenbrunner et al. 2009)]
narratives emerge could be made up of 100 posts, 99 of which support a given view, and 1 of which offers an alternative view. Even if this one lonely post is excessive in length, it can hardly be overlooked that the views of one side to the debate are overrepresented – but this piece of information we might lose if we focused on threads as units of discussion.

So I take individual comment entries (posts) as units of content analysis. It needs to be noted that breaking the discussions down to their individual building blocks results in losing certain pieces of contextual organization (e.g. we might become aware that a person is a fanatic supporter of a particular party, which might influence our interpretation of this person's messages which do not in themselves contain any information about their author's party affiliation).

6.3. Coding categories for the content analysis

6.3.1. Heterogeneity of expressed views

As explained in the previous chapter, I will use as a first measure of heterogeneity the ratio of pro-Democrat versus pro-Republican versus pro-third party messages. That is, each comment entry in a sample is to be coded either Democrat, Republican or third party, and the ratio of these messages to each other are to be used as a measure of heterogeneity of views.

The classification of messages is to be decided by asking a question: inferring from the text of the comment (viewed in separation!), which party did its author support in the 2008 presidential election?

If the author's party preference cannot be decided with great certainty, then the entry is to be coded uncertain.

In deciding the party preference of the author, I appeal to two simplifying assumptions. First, negative attitude towards one of the major parties should be interpreted as supportive attitude towards the other major party, unless otherwise stated by the author. For instance, a comment dissing John McCain as a candidate is to be coded as supporting the Democratic party, even though the author might not say anything about Obama, Biden or the Democrats in general. (For the rationale of this assumption, see Ashbee 2004: 184 – 188 (and my paper on the perceived bias on social news sites).)

And second, opinion on certain issues is to be used as an indicator of party preference. For example, taking a so-called “pro-life” position in a debate about abortion is to be interpreted as support for the Republican party, and arguing for the fastest possible military exit from Iraq is to be interpreted as support for the Democrats.

This scheme is to provide at least an approximative picture of the ideological variety and balance expressed in debates on social news sites.

6.3.2. Argumentativeness and rationality

As I wrote in chapter 5, online discussions can only be considered deliberation provided that they contain (a) arguments which are (b) valid.

(a) For the definition of argument, I turn to the general scheme of argumentation developed by Toulmin, Rieke and Janik (1979). In their approach, an argument is a claim supported by some grounds, e.g. “it can be stated that [claim], because [grounds].”

A structurally complete argument also includes a warrant – explaining the connection between the grounds and the claim –, and a backing or evidence of the warrant. In addition, the argument might also include some modal qualifier (such as “presumably, likely, certainly...”), and some specific rebuttal or condition which would render the argument invalid. Thus the abstract structure of a
complete argument could be described as follows: given certain *grounds*, we may appeal to *warrant* (which rests on a particular *backing*) to justify our *claim*, or at any rate the *preumption-that-claim*, in the absence of some specific *rebuttal* or disqualification.

Thus the sentence “We should vote for [John] McCain, because he has more experience [than Obama]” should be understood as an argument, making the claim that we should vote for McCain, on the grounds that he has more experience. The implied warrant in this case is the idea that experience is a relevant indicator of one’s presidential capabilities. Similarly, a somewhat extended argument would be “We should presumably [modal qualifier] vote for McCain [claim], because he has more executive experience [grounds], unless he is found to be in too frail physical condition [contingent rebuttal].”

However, not all arguments will be considered *relevant* in our analysis. In order to find out which ones will be relevant, again it is necessary to refer to the context of the analysis. The problem that the analysed deliberations should focus on, is simple in abstract terms: from a small number of candidates, one should be chosen as the next president of the country. Therefore those arguments should be counted as relevant which address this problem by suggesting that one or another candidate is to be selected (or, in the contrary, that selection of one or another candidate is to be avoided). Appealing to the same simplifying assumption as in the case of ideological heterogeneity and perceived bias on social news sites, I then propose that only those arguments should be retained in the analysis as relevant which make the claim that a particular candidate (or a particular party) is to be supported in the 2008 US presidential election. Such a claim need not be made explicit; it also can be implied by the author of a comment message. So, a comment post contains an argument if it contains the equivalent of the sentence “[grounds], therefore we should vote for X (party or candidate) [claim].” These kinds of argument I will refer to as *direct* arguments. On the other hand, there is another kind of relevant argument that can be made: arguments which do not directly address the question of the debate (who to vote for), but which address a previous argument (or a previous expression of views on who to vote for). These arguments I will call, conveniently, *indirect* arguments. For example:

A: “I think we should vote for McCain [claim], because he was a prisoner of war [grounds].”

B [in reply to A]: “That he was a prisoner of war has nothing to do with his qualities as a president.”
In this example, B opines that the warrant that is implied in the argument made by A is faulty, or at least questionable. B might, or might not, support McCain – we don't know that. His response doesn't make a claim of whom to support, it merely makes the indirect claim that a particular line of reasoning is incorrect.

It is trivial that indirect arguments, when standing on valid grounds, are highly relevant to the deliberation, even though they themselves might not propose a solution to the problem debated.

(b) For an argument to be considered valid, it needs to be relevant, and it needs to be rational. But how should rationality be understood?

As Habermas (2005: 391) notes, different, competing concepts of rationality exist. It is in his *Theory of Communicative Action* (1984, vol. I: 8 – 43) that he lays down the foundations of his theory of rationality that continue to influence his model of ideal deliberation. Habermas distinguishes between instrumental rationality (the deployment of knowledge in teleological action), and communicative rationality. This latter means the “employment of propositional knowledge in assertions” in order to achieve intersubjectivity.

In other words, communicative rationality means aiming at understanding others and making ourselves understood; it is concerned with appealing to such objective, external standards, by which our individual, personal thoughts, ideas and feelings can be transmitted to and interpreted by others. This idea of communicative rationality can be translated into practice by considering those arguments rational whose grounds appeal (a) to objective, external facts that can be empirically confirmed or denied, (b) to shared norms or (c) to authentic subjective feelings and experiences.

The first of these categories is self-explanatory: the simplest form of rationality ties the argument to an objectively and independently existing fact. But Habermas argues that shared norms (the acceptance of which ultimately depends on certain values that have no rationally explainable factual grounding), and personal feelings, emotions and experiences can also serve as “good reason” for accepting a particular argument. It is perfectly rational for me to argue for a course of action on the grounds that it would be right because of certain values; or on the grounds that I experience a particular emotion

I believe the important thing here is not to confuse the validity of arguments with their quality. The above definition of rationality is very broad. It extends to arguments that in everyday thinking would be immediately judged irrational. Consider the statement: “Obama is white, so we should vote for him as president.” This is an obviously incorrect argument, insofar as it rests on grounds that are objectively untrue, and leaves the connection between the grounds and the claim unclarified (i.e. even if we supposed that Obama was white, that would still not explain why we should vote for him). But it is a rational argument, because it offers external, factual evidence in its support, which can be externally verified.

So if we do accept the above definition of rationality, then it becomes trivial that just because an argument is rational, that does not mean that it is necessarily a correct or good argument. A rational argument can also be incorrect; and correct and rational arguments can also be weak, in the sense that they would not be generalizable to other members of the group that is going to be affected by the eventual decision after the debate.

Deliberation requires arguments to be valid – because otherwise no debate could take place –, but the point of the debate is to find out which of these valid arguments are better than the others; to

---

4 Indirect arguments are not to be confused with *ad hominem attacks*, i.e. arguments that are aimed not against a previous argument, but against the *maker* of a previous argument, often on personal grounds.

5 This definition of rationality is so broad that one might wonder, how could an argument be anything but rational. But an argument would certainly not be rational if it omitted its grounds, or phrased them incomprehensibly, e.g. “All mimsy were borogoves”, hence McCain should be voted president. (cf. *Jabberwocky* by Lewis Carroll, in *Through the Looking-Glass and What Alice Found There.*)
find out which arguments could be accepted by the largest possible group of people. This is not to say that deliberation is worse off if weak or incorrect arguments are presented. On the contrary: proposing rational, but weak or incorrect arguments provides other participants with an opportunity to point out the flaws in these arguments. And if the conversation is also reflexive, then this means that those who made the weak arguments will forfeit them and look for better ones.

In my coding scheme, I attempt to classify the relevant and valid arguments according to the grounds they appeal to, as I suppose that certain grounds of arguments are more likely to result in better arguments than others. Thus the kinds of grounds appealed to can serve as an indicator of the quality of arguments used. The classification I employ originates in the particular context I study deliberation in.

Deliberation in my case is concerned with the problem of choosing the candidate that is expected to be the best president. This problem refers to the future: in no way can we ascertain before the elections that the candidate eventually chosen will take the best course of action. Our opinions cannot be based on future facts; so we have to turn to the past to offer some guidance as to what we can reasonably expect from the candidates. With this in mind, I divided the range of possible grounds (or reasons) of direct arguments into three categories.

(1) **Grounds related to policies.** Such grounds state that particular policies or political values are better than others, and that a given candidate, based on his or her program, voting record or other evidence, is likely to execute these particular policies. If we understood the concept of rationality in a much stricter sense than Habermas, then arguably this kind of arguments would be the only one we should count as relevant: our society faces certain problems, which need to be addressed by policies, so it is rational and reasonable to make up our minds about the candidates solely on the basis of the policies that they advocate.

(2) **Grounds related to the personal life or qualities of the candidates.** The problem of electing a president can also be translated as follows. We need to decide who is going to make the best decisions as president. One way to see who the best decision maker would be is to judge them on the basis of their perceived personal attributes. One might argue that the solutions offered by candidates to the current set of societal problems might not indicate correctly what their solutions would be to new, as yet unknown problems (which eventually will surely present themselves). So it might be that their personal attributes are a better indicator of their overall decision making skills. While admitting to the importance of candidates' personal qualities, I consider arguments falling into this second category as potentially weaker than policy-based arguments, mainly because of the uncertainty involved in both the perception and in the evaluation of personal qualities (which problems do not, at least in theory, exist in connection with policies and political programs).

(3) **“Other” grounds.** To the third category belong arguments that are based neither on the policies, nor the personalities of the candidates.

Each argument is to be classified as belonging to only one of the above categories. Naturally, a single comment post may contain any number of different arguments – each of which are analysed and coded separately. Doing so will allow me to reach a quantitative analysis of arguments within a sample of conversation.

---

6 E.g. the sentence: “Obama should be president, because his tax policy is solid (as opposed to that of McCain), and he is also pro-choice” contains two arguments. The claim is the same in both of them – Obama should be voted for –, and in the first argument, this claim is based on grounds related to his tax policy, in the second argument, it is based on grounds related to his views on abortion. (Two policy-related grounds.)
Finally, it is important to note that the dimension of justification and rationality is independent from the dimension of expressed ideology of a comment post. A comment post can be coded as “uncertain” as for its expressed ideology – yet it may well contain relevant and valid arguments. And it also might happen that a post contains arguments that seem to contradict the ideology expressed in it; for example: “I also support McCain, but it has to be noted that Obama's economic policies are better formulated” (the argument supports Obama, but the author underlines that in spite of this – for reasons not mentioned in this sentence – he does not suggest that he should be voted for).

6.3.3. Deeper into the arguments

The format of online discussions presents an opportunity to analyse not only the presence and contents of arguments, but also the evidence offered in support of their grounds. Not only is online communication written, it also provides the option of hyperlinking to relevant pieces of information. If an argument is made using some external facts as grounds, evidence of the truth of said facts can be conveniently provided in the form of a hyperlink reference. Much like references in scientific works, this practice is bound to increase the transparency of discussions, and to increase the force of arguments. It is harder to dismiss an unlikely argument if the ground it stands on is demonstrated to be true. The concept of evidence for grounds does not feature the general model of argumentation by Toulmin, Rieke and Janik. (The concept of “backing” refers to evidence of the “warrant,” i.e. the connection between the grounds and the claims of the argument.) But it can be easily inserted into the model: On particular grounds I make a claim – and I cite evidence that shows that the mentioned grounds are true. (Figure 4.)

![Figure 4. An extended model of arguments. (Cf. Toulmin et al. 1979: 78)](image)

In the coding process, coders are to decide whether arguments include such verifiable, external evidence or not. (By the nature of the concept, evidence needs to stated explicitly.) A self-explanatory way of providing such evidence is through the use of hyperlinks. This however is not the only possible way; for instance, an unambiguous reference to a book or newspaper article would also count as valid evidence.

The quantitative analysis of the evidence offered in support of grounds is another aspect of measuring how the observed online discussions compare to a theoretical ideal deliberation.
6.4. Measurement methods – summed up

In order for online discussion to qualify as deliberation – in the particular context I'm focusing on –, it needs to fulfill the following criteria: it needs to present heterogeneous views, it needs to be externally and internally open, it needs to be based on valid arguments, it needs to be reciprocal, and participants to it need to exercise reflexivity (chapter 5).

I suggest that three of these features be analysed from the written, archived text of the conversations themselves; namely the heterogeneity of views, the presence and quality of arguments used, and the reciprocity of the discussion.

Reciprocity is to be measured through the structural analysis of conversation threads. The heterogeneity of represented views, and the arguments used in conversation are to be studied through the analysis of the contents of the discussions. Individual, separate comment posts are to be treated as units of analysis.

As for the heterogeneity of views, coders are to code comment entries as supporting either the Democratic, the Republican, or a third party; or as uncertain, if no strong signs of ideological affiliation can be detected.

An argument is understood as a claim backed up by some grounds. Relevant arguments are those which claim (explicitly or implicitly) that a particular candidate should be voted for in the election (direct arguments), and those which subject a previous argument to criticism (indirect arguments). Arguments are considered valid if they are rational, that is, if they provide grounds that appeal to externally verifiable facts, shared norms, or an authentic personal emotion or experience. Valid, relevant and direct arguments are to be classified according to their subject matter as arguments related to policies, arguments related to the personality of candidates or other arguments. All arguments are also to be classified in function of the evidence they offer.

The proposed coding scheme does not attempt to measure the openness or reflexivity of online discussions. But while incomplete, I believe it could be indicative of the deliberative potential of online discussions on its own; and naturally it could subsequently complemented with the analysis of the remaining characteristics of discussion.

The development of standards for the interpretation of the analytical results remains an important task of the research project.

7. A practical example

What follows is a demonstration of the coding scheme developed above, through the analysis of a discussion sample taken from the social news website Reddit.

7.1. The sample

The news item that is in the focus of discussion is a video clip from Comedy Central's “The Daily Show with Jon Stewart” programme, aired on 5th September 2008. It was submitted to Reddit by user “bgholt1970” on 8th September 2008, and on that day it was the most popular campaign-related item in the contents of Reddit. The clip – roughly 9 minutes long – makes fun of John McCain's acceptance speech at the Republican Party's national convention (where he officially became the presidential candidate of the GOP). Notably, The Daily Show's clip compares the speech of McCain to the acceptance speech of George W. Bush from 2000, revealing how similar the messages are, often repeating the same claims verbatim.

The originally submitted item pointed to the liberal “The Information Paradox” blog, from where it was removed. However, the first commenter posted a link to the clip as it is featured on the
homepage of Comedy Central itself. The clip was submitted under the title “Jon Stewart Obliterates McCain's Acceptance Speech”. It received 1263 “points” on Reddit, meaning it received 1263 more “up” than “down” votes.

7.2. Reciprocity

The item received 246 comments, 11 of which were subsequently deleted (either by the decision of their authors, or by moderators of the site). Thus the discussion spans 235 individual comment posts.

In the discussion, 114 sub-threads can be distinguished (conversation width: 114). Consequently, the average depth of these sub-threads is 2.06. A more accurate picture is provided by the separate count of the depth of sub-threads:

<table>
<thead>
<tr>
<th>Depth</th>
<th>Number of sub-threads</th>
<th>Share of sub-threads (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>10.53</td>
</tr>
<tr>
<td>2 (mode)</td>
<td>26</td>
<td>22.81</td>
</tr>
<tr>
<td>3</td>
<td>22</td>
<td>19.3</td>
</tr>
<tr>
<td>4</td>
<td>17</td>
<td>14.91</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>13.16</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>6.14</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>7.02</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>5.26</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0.88</td>
</tr>
<tr>
<td>Total</td>
<td>114</td>
<td>100</td>
</tr>
</tbody>
</table>

As this table testifies, the mode (the most frequent value of depth) was 2, very close to the average. Is this a good result?
A 2-reply-deep thread suggests that someone made a comment on the original news item, to which someone else replied – but here the conversation stopped. If there was a counter-argument offered to the first comment, it is likely that it was ignored. Arguably, a minimal requirement of reciprocal deliberation would be a conversation at least 3 replies deep: “A” makes an argument, to which “B” replies, to which “A” replies, presumably with a counter-argument or with conceding to the points made by “B.”

On the other hand, exactly two thirds (66.66%) of the sub-threads are in fact 3 or more replies deep; and as the table testifies, in at least a few cases such discussions developed which provided ample opportunity for the participants to elaborate on their points of view.

7.3. Content analysis

For reasons of convenience, the sample of this demonstrative content analysis includes the first 75 comments from the chronologically ordered discussion thread.

7.3.1. Heterogeneity of views

The following table sums up the variety of ideologies represented in the discussion:
An ideal scenario need not require that the Democratic, Republican and other, third parties are represented to an equal extent in the conversations – this would perhaps be a too shallow interpretation of the concept of heterogeneity. What is notable, however, is that the majority of comments in the sample was coded “uncertain” – in many cases for the reason that the post had nothing to do with the topic of election, or at least it only connected to the topic in very loosely. (E.g. a lengthy conversation developed about the problem of watching streamed online videos using proxy servers.)

Also notable is the absence of third party-ideologies from the discussion, and the threefold overrepresentation of views supporting the Democratic party (which result, by the way, is consistent with the ideological bias found in the published content of Reddit).

Since the point of this paper is to demonstrate the coding scheme in operation, I cite and analyse here some examples.

- **a.) Comment**: “Now without blogspam! Voted up parent; voted down topic.”
  **Interpretation**: This comment post was a reply to another comment, in which the author provided a direct link to the homepage of The Daily Show, where the clip was accessible. “Now without blogspam!” refers to the fact that the poster disliked The Information Paradox blog (the source where the submitted video clip was hosted). “Voted up parent; voted down topic.” means that the poster voted positively for the “parent comment” (the one he replied to), and negatively for the discussion topic itself.
  **Judgement**: the post was coded as supporting the Democratic party, since it gave explicit support (voted up parent) to content which unmistakably criticized and made fun of the Republican Party in general and John McCain in particular. (Conversely, another comment which consisted of “Bias alert! Voted down” was coded as supportive of the Republicans.)

- **b.) Comment**: “latest gallup poll gives Obama only 2% edge over McCain – a statistical dead heat again”
  **Judgement**: the post was coded as “uncertain”, as for its expressed ideology.

- **c.) Comment**: [in reply to b.)] “That's just so sad, I mean, isnt there something like 80% who think the last 8 years haven't been good... But ~50% still want republican leadership?”
  **Judgement**: the post was coded as supporting the Democratic Party; expressing disbelief and disapproval of the apparent behaviour of voters. (“How is it even possible that people still want Republicans in the White House?!”)

- **d.) Comment**: [in reply to c.)] “I think the last administration sucked, but I don't blame Republicans. The democrats took over the house and senate over the past, what, 2+ years, and nothing has changed, has it? People act like Bush controls the universe. Fact is, Republicans want me to take care of myself. Democrats support, simply put, the redistribution of wealth. I do not agree with the socialist agenda of the Dems.”
**Judgement:** the post was coded as supporting the Republican Party. Although it does not explicitly say that the Republicans exercise good policies, it expresses disapproval of Democratic policies, and the “socialist” idea of redistribution of wealth.

### 7.3.2. Argumentativeness

The sampled 75 comment posts contained 19 relevant arguments (suggesting or implying that a particular candidate is to be voted for on certain grounds); 8 of which was coded as direct (related to the candidates or their respective parties), and 11 as indirect (subjecting a previous line of argument to critique). The ratio of arguments to number of posts (19/75) is 0.253, which could be translated as every fourth post containing an argument – far from what would be expected from ideal deliberation.

An example of direct and indirect arguments:

-e.) **Comment:** “[...] Fact is, Republicans want me to take care of myself. Democrats support, simply put, the redistribution of wealth. I do not agree with the socialist agenda of the Dems.”

**Judgement:** direct, policy-related argument.

-f.) **Comment:** [in reply to a.)] “I think the last administration sucked, but I don't blame Republicans. The democrats took over the house and senate over the past, what, 2+ years, and nothing has changed, has it? People act like Bush controls the universe. [...]”

**Interpretation:** That people are dissatisfied with the previous administration, and still support McCain is not a paradox. It is incorrect to blame the problems of the past 8 years on the GOP and on Bush in person. Such an argument would be wrong.

**Judgement:** indirect argument.

The further analysis of direct arguments reveals that all of them related to policies or issues; and consequently none of them offered grounds related to the personalities of the candidates, or to some other factor. What follows is a list of comments that contained the arguments in question, and the interpretation of their grounds. (As you can see, the 8 direct arguments offered were contained in only 5 posts.)
<table>
<thead>
<tr>
<th>Comment</th>
<th>Claim (what shall we do?)</th>
<th>Grounds (...why?)</th>
</tr>
</thead>
<tbody>
<tr>
<td>“[...] I don't agree with the socialist agenda of the dems.”</td>
<td>We shall vote for the Republican Party's candidate (=we shall not vote for the Democratic Party's candidate.)</td>
<td>Because the socialist agenda – the idea of redistribution of wealth – is wrong.</td>
</tr>
<tr>
<td>“Under Democrat policy &amp; Clinton, things were never this bad. Under Republican policy, things are as horrible as I have ever seen. Edit: it isn't just about the super rich getting richer and the middle class getting poorer [1]. Everything contributed. From their aggressive military posturing [2] to their failed foreign policy [3]. As well as (lack of) funding of important public work projects.[4]”</td>
<td>We shall vote for the Democratic Party's candidate.</td>
<td>[1] Because under Republican policy inequalities of wealth have substantially grown. [2] Because Republican policies lead to aggressive military posturing. [3] Because Republican foreign policy failed. [4] Because Republicans did not provide funding for important public work projects.</td>
</tr>
<tr>
<td>“Yes Republicans do want you to take care of yourself, while they and their friends get rich. Capitalism just like socialism works, but it can be wrong just as easy.”</td>
<td>We shall vote for the Democratic Party's candidate.</td>
<td>Because the Republican idea of capitalism is wrong (it exploits the poor).</td>
</tr>
<tr>
<td>“Republicans want to meddle more with your personal life (abortions, same-sex marriage and all the other religious issues) than democrats.”</td>
<td>We shall vote for the Democratic Party's candidate.</td>
<td>Because Republicans want to meddle with your personal life more than Democrats (and “meddling” with your personal life is wrong). [Cf. the negative connotation of the verb “meddle”.]</td>
</tr>
<tr>
<td>[in reply to: “The Republicans want you to take care of yourself.”] “No, the republicans want you to take care of rich people.”</td>
<td>We shall vote for the Democratic Party's candidate.</td>
<td>Because the Republican idea of capitalism is wrong and exploitative.</td>
</tr>
</tbody>
</table>

Finally, the analysis of evidence offered in the support of arguments reveals an almost complete lack thereof. Out of 19 identified arguments, only 2 were supported by external evidence – both of them being indirect arguments.

In the first occasion, the debate revolved around the question of whose fault the past 8 years were. Apparent Obama supporters argued that only the Republican Party should be blamed, because even though the Democrats secured a majority in the latter half of the 2nd Bush term in the House of Representatives and the Senate, these majorities were not enough to change the course of action greatly, because Bush used his presidential power to veto important pieces of legislation. In this debate, an apparent Republican supporter claimed that “Bush has hardly ever used the veto, far less then [sic] almost every other president.” In response to this claim, a commenter posted an exhaustive list of the 12 acts Bush vetoed, providing an external hyperlink to a Wikipedia page containing a list of vetoes by all the presidents of the USA.

The second occasion introduced an interesting problem. A commenter posted the hyperlink of a pollster, suggesting that it shows “a very likely Obama victory and a possible blow-out.” After some hesitation, this post was coded as supporting the Democratic Party, largely on account of the expression “blow-out.” On the other hand, the post did not contain any argument, relevant or irrelevant; therefore the external evidence it offered in support of its claims did not gain significance.

However, an apparent Republican supporter replied to this post, claiming that “That's not a blowout.
It's a two state swing with a lot of weak support and undecided states.” This post was coded as containing an indirect argument: it referred to a previous line of thought and pointed out a flaw of interpretation in it. Even though the comment this indirect argument referred to was itself coded as non-argument, I felt the classification of indirect argument to be justified. And since this second argument did refer to an externally hyperlinked, verifiable source, it was coded as citing external evidence.

7.4. Summary

This exploratory analysis revealed that the sampled discussion largely failed to live up to deliberative ideals; although there are uncertainties concerning the interpretation of the findings.

As for the reciprocity of the discussion, the width of the discussion was found to be 114, while its average depth amounted to 2.06. On the other hand, 66% of the distinct threads within the conversation ran at least 3 replies deep. This might be an important measure if we suppose that deliberation should ideally incorporate at least 3 subsequent replies.

As for the heterogeneity of views expressed in the discussion, it was found that a.) the Democratic Party's ideology was overrepresented, b.) no comment expressed support of a minor, third party, and c.) the majority of comments were coded as uncertain in their expressed support, possibly because they contained off-topic messages.

As for the argumentativeness of conversation, it was found that on average, only 1/4th of the comments contained relevant, either direct or indirect, arguments. The majority of arguments were indirect (referring not to the candidates or their parties, but to the points raised by other commenters), but all of the direct arguments registered related to policies (as opposed to the personality of candidates or to some other factor). However, evidence was only offered for two arguments (both of which was coded as indirect).
Works cited


Cambridge: Cambridge University Press.


