

Managing Conflict in Online Debate Communities: Foregrounding Moderators' Beliefs and Values on Kialo

Jordan Beck, Bikalpa Neupane, & John M. Carroll
The Pennsylvania State University, University Park, PA

Abstract

Kialo is a novel peer production system focused on pro/con debate construction. Teams of moderators vet and accept claims submitted by writers. Moderators also edit and refactor debates as they grow. Thus, moderators play a critical role in cultivating and maintaining debates. Conflict between moderators is typical. It is a feature of argumentation and debate. However, not all conflict is productive. Conflict between moderators can undermine collaboration (by distracting from the task of managing debates) and drive attrition (by discouraging participation on the site altogether). Based on a ten-month participant observation on Kialo, we identify a common source of conflict between moderators: adversarial beliefs and values. Moderators are not neutral participants on Kialo. They take positions on debate topics. We suggest foregrounding these positions, which are potential sources of conflict, through interface design as a scalable solution to conflict management.

Keywords: peer production, debate, debate moderation, kialo, conflict.

1. Introduction

Kialo (www.kialo.com) is a novel online debate platform that can be understood as a community of practice (Wenger 1998). Its participants cooperate to develop complex pro/con debates about a variety of topics, including: net neutrality, ad blockers, artificial general intelligence, climate change, gender, and reproductive rights, among others. A given debate may involve hundreds of participants and thousands of pro or con claims. Debate participants—regardless of their role—engage in continual learning processes about (1) a topic (e.g., net neutrality), (2) the theory and practice of argumentation, and (3) how to use Kialo. Admins and editors, which we refer to in this paper as “moderators,” play a central role in the development of a debate. Moderators make decisions about who gets to participate, which claims are accept-

able, as well as how to revise claims in order to strengthen their overall contribution. As with other researchers studying peer production systems (Luther et al. 2013; Kim et al. 2014) we believe that the moderator role is vital to the growth and success of particular debates.

Moderators on Kialo guide other participants’ learning; especially with regard to how to make good claims. For instance, moderators critique suggested claims for their lack of clarity or lack of support. They also monitor debates for duplicate claims, insincere contributions, and vulgar/abusive content. Moderators work in teams, which means that collaboration is a key of their practice. Few (if any) actions a moderator can take in a debate are carried out in isolation. For example, moderators discuss suggested claims with each other before deciding to accept them in a debate. This is a kind of collaborative “gatekeeping” (Keegan and Gergle 2010), and it is ultimately visible to the Kialo community and the public. Moderators also discuss accepted claims, i.e. whether they need supporting evidence or lack clarity, and/or whether they remain relevant as a debate changes over time. It is common that, as a debate grows, the framing or main thesis may change, which motivates the moderator team to reevaluate and refactor the entire debate.

Kialo provides moderators with a set of tools that enable collaboration. These tools shape the ways moderators interact with each other. For example, moderators can “flag” problematic claims for a finite number of predetermined reasons, they can provide feedback to writers, and they can discuss problematic claims with other moderators and writers via two separate chat tools. However, the Kialo toolset changes with some regularity.

In some cases, these changes create conflicts to arise between moderators and inspire moderators to re-negotiate the norms and conventions of their practice. On the basis of an ongoing (ten months) participant observation, in this paper, we discuss one such interface change and the resulting conflicts and negotiations between moderators. As it turns out, debate moderation on Kialo is not value-free. Moderators have stakes in the debates they moderate, and these stakes influence their decisions and actions as moderators. We argue that there is utility and value in knowing what these stakes are and propose that Kialo—and other peer production communities—make personal stakes more visible so that other members of these communities can leverage this knowledge for collaboration.

2. Background

Moderators play various roles in different online communities. For example, they can serve as defenders against trolling and flaming on discussion boards. They can serve as “project managers” assigning tasks to participants producing animated movies (Luther et al. 2013) or music albums (Settles and Dow 2013). Moderation has been studied in peer production communities (Zhu et al. 2011, 2012; Arazy et al. 2015), creative collaborations (Kim et al. 2014), and on online news sites (Park et al. 2016). More recently, there have been opportunities to examine moderation in the context of online debate and deliberation systems (Kriplean et al. 2011, 2014). A particular interest has been the roles moderators play in reducing barriers to participation, including conflict and dispute resolution.

Moderation in Peer Production Communities

Moderators can perform a variety of roles in online peer production communities, which are distinguished by facilitating and coordinating the work of a large number of people toward a

shared outcome. Common examples of successful peer production communities are Wikipedia and Linux. In these communities, moderators can be responsible for managing participants (Krieger et al. 2009), crafting creative or intellectual project visions, helping maintain quality standards (Liu and Ram 2011), and “gatekeeping” (Keegan and Gergle 2010) against vulgar or abusive participants. Regardless of the community under examination, existing studies of online moderation share an assumption that moderators are crucial to online peer production and attempt to do one of two things: (1) Theorize online moderation by identifying barriers and analyzing cases and (2) propose policies or technical solutions to make online moderation more effective.

Challenges to Effective Moderation

Researchers have identified a number of challenges that undermine moderation efficacy. Some are related to the number and type of tasks. For example, Luther, Fiesler, & Bruckman (Luther et al. 2013) found that moderators in a flash animation community became overburdened with too many tasks and responsibilities, which caused projects to stall and fail. Providing timely, quality feedback to participants can also be challenging (Dow et al. 2012), especially given that most moderation is done on a volunteer basis. This means that moderating competes with other personal and professional responsibilities.

Conflict is another key challenge moderators face. Peer production systems, such as Wikipedia, involve interaction between humans, which means that conflict is bound to happen. Moderators may encounter conflict between participants or with other moderators, and it can lead to a few possible outcomes. Conflict can be detrimental to peer production communities by causing projects to “stall or fail” (Billings and Watts 2010), by discouraging participation (Huang et al. 2016), and by leading to the production of low-quality content. On the other hand, conflict can be productive (Kittur et al. 2009) and useful. For example, it can enable people to challenge their own perspectives on a social or political issue (Kriplean et al. 2012). Moreover, when it comes to peer production, conflict (in the form of debate or argumentation) can result in higher quality outputs. For example, Arazy, Yeo, and Nov have studied the relationship between debates between editors and the quality of Wikipedia articles (Arazy et al. 2013). These possible outcomes, and others, have led researchers to develop an interest in understanding online conflict and conflict management (Filippova and Cho 2016; Fréard et al. 2010; Grevet et al. 2014).

Understanding and Addressing Online Conflict

Researchers have proposed different factors that contribute to online conflicts. Some have argued, for example, that task interdependence and geographical distribution in some cases increase conflict in free and open-source software development teams. Schneider et al. (Schneider et al. 2013) found that arguing with Wikipedia collaborators on the basis of “personal preference and inappropriate analogy to other cases,” rather than adhering to community norms and conventions, can be seen by others in the community as problematic, and, thus, fuel conflicts. In addition, disagreements among leaders (e.g. debate moderators) about processes and procedures can be interpreted as, or become, personal, which can distract from the tasks at hand. Conflicts in open collaboration and peer production communities also arise due to malleable or poorly-defined policies and/or ideological issues (Filippova and Cho 2015).

Resolving or managing online conflict in these settings can be consequential both for the work

being done (e.g. authoring an article or constructing an argument) and for the general well being of the community (e.g. people enjoy participating).

Scalable Conflict Management Strategies.

Different proposals for managing conflict have been put forth, and some have been deployed. On Wikipedia, some participants work as “mediators... [helping] conflicting parties to express, recognize, and respond positively to their personal and substantive differences” (Billings and Watts 2010, p. 1447). Others have proposed that responding to other contributors/collaborators with constructive suggestions for improvement is more effective than rational explanations of problems or generic social encouragement when it comes to managing conflict (Huang et al. 2016). Finally, utilizing participatory decision-making and certain leadership styles have been shown to mitigate conflict in free and open source software development communities (Filippova and Cho 2016). There are few examples of conflict management in online argumentation systems. However, we interpret Kriplean et al.’s work in (Kriplean et al. 2014) as an indirect example of conflict management.

Kriplean et al. developed *ConsiderIt* to support public deliberation. Users work together to create pro/con debates about pertinent local social and political issues. Citizens contribute claims about real issues, such as the first-ever Washington state income tax (Kriplean et al. 2012, p. 267). The researchers recognized the importance of having reliable information about these kinds of issues. For example, if a user claimed, “The state legislature may expand the income tax to the middle class in two years,” (Kriplean et al. 2012, p. 267) it would be important to vet this claim. At the same time, leaving the vetting up to other users could produce contentious arguments (conflict). Anticipating the possibility of these conflicts, the researchers recognized the need to involve moderators who would be both reliable and seemingly neutral. So, they recruited public librarians to work as on-demand fact-checkers (Kriplean et al. 2014). Users responded favorably to the librarians’ role even when they disagreed with the fact-check.

This is a promising outcome, but the authors draw attention to the issue of scalability. Just as other volunteer moderators are pressed for time, so too were the public librarians. A key question becomes how to devise strategies for conflict management and resolution that leverage interface and interaction design?

2.1. Summary

Although there are a number of peer production systems that exist to support argumentation, to our knowledge, none have been studied directly in terms of conflict between participants. However, it is also clear that conflict between participants might be of significant interest. Kriplean et al.(2012), for example, describe normatively desirable activities on ConsiderIt to include “crafting positions that recognize pros and cons as well as points written by people who do not agree with them” (Kriplean et al. 2012, p. 1). That is, it is desirable for users to manage conflict such that they engage with disagreeable points of view. Similarly, Widescope (Burbank et al. 2011) aimed to facilitate dialogue between people with divergent points of view with the goal of achieving some objective, such as arriving “at a mutually acceptable compromise.” Managing or resolving conflicts is crucial to achieving such an objective.

Second, of the existing solutions to conflict management and resolution in peer production

systems, most tend involve policy prescriptions for how human actors ought to behave towards one another. For example, providing constructive suggestions, using language that adheres to the norms and conventions of an online community, and developing special roles (e.g. mediators) for participants to adopt and perform, are descriptions of human actions. These are good and reasonable suggestions. In this paper, we propose that it is also important to explore ways that peer production systems might be designed to support or augment human actions.

3. Kialo

Given that Kialo is a novel peer production community, in this section, we briefly summarize the structure of debates and the different roles participants might play. In particular, we describe how people suggest new claims, the importance of evaluating those claims, and the ways moderators conduct said evaluations.

Kialo uses the structure of dialectical reasoning to explore different sides of an issue. Each debate has a main thesis, or several main theses, which are elaborated through ‘pro’ and ‘con’ claims. Anyone with a Kialo account can start a debate on any topic. To our knowledge, there are no restrictions on what topics are up for debate, and, in fact, there is a somewhat burdensome process to go through to delete a public debate from the site. Debates can be public or private. Private debates are invite-only whereas public debates are in principle visible to anyone with an Internet connection, the knowledge that Kialo exists, and the time and interest to search its growing set of debates.

Kialo has established a set of roles that participants can play in a given debate. These are:

Admin: Admins can modify discussion settings, change the rights of users and invite new members. Admins control whether a discussion is private or public, can change the discussion between single- and multi-thesis forms, write tags, and change the cover image. They are also able to accept suggested claims or mark them for review.

Editor: Editors have full rights to create, edit, and delete [all] claims [in a debate], as well as to comment or mark claims for review.

Writer: Writers have the rights to create, edit, and delete [their own] claims, as well as to comment or mark [all] claims for review.

Viewer: Viewers can see all the content in a debate, and they do not need a Kialo account. They have no rights to do anything in a debate except view its content, [which includes any/all communication between active participants in a debate].

Viewers can suggest claims. However, they have to wait for moderators—admins or editors in Kialo’s terms—to vet and accept their suggestions. Once a person has been granted Writer status, they can add claims without a priori vetting. Vetting claims seems to be one of the most important parts of moderating a debate. As one moderator explained to us, “*Badly worded claims... invite more badly worded claims*” @libre. Vetting claims is thus seen as directly contributing to the overall quality of a debate.

When we started our participant observation in October 2017, vetting suggested claims was an individual process. Moderators could accept or “send back” a claim without consulting

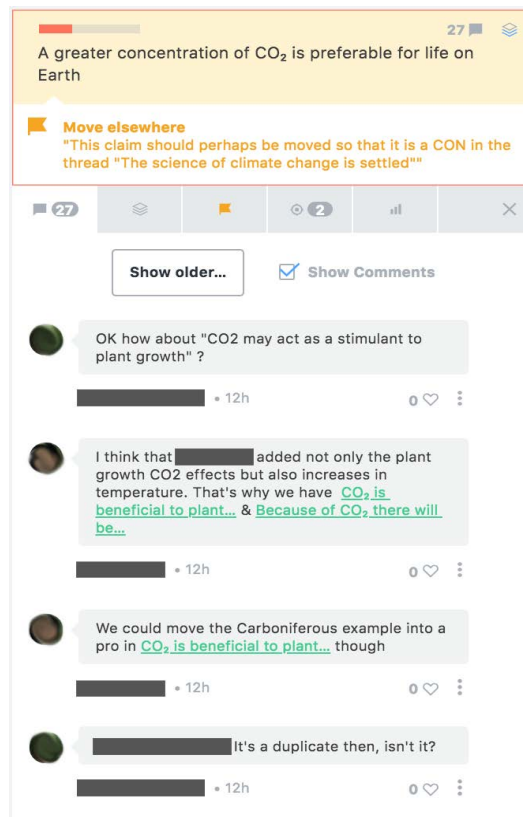


Figure 1: Moderators Discuss a Flagged Claim on Kialo.

other moderators. Sending claims back to a writer was the only aspect of moderation that was hidden from the broader community. Only after a claim had been accepted, if a moderator took any action (e.g. flagging or commenting on it), then that action would be publicly visible. In early 2018, however, Kialo implemented a change that kept suggested claims static and visible to all moderators, which turned suggested claim vetting into a collaborative activity (Figure 1). It was no longer possible to send claims back. Instead, a moderator could 'reply' to the claim with comments, questions, or revision requests, all of which would be visible to the moderator team.

4. Methods

In line with existing studies of online communities (Boellstorff et al. 2012), we made the decision to use virtual ethnographic methods to explore moderation practices on Kialo. We are engaged in ongoing (ten months) participant observation, which means that we are active writers and moderators in several debates on Kialo.

Data Collection

We have been developing a thick record (Carspecken 1996) of our interactions on Kialo. This consists of (1) low-inference summaries of interactions and experiences on Kialo and (2) relevant, publicly visible user-generated text on Kialo. This publicly visible text comes either from a discussion chat or a claim chat. The discussion chat facilitates talk about high-level issues pertinent to a debate (e.g. are there too many top-level claims, does the main thesis need to change), onboarding new participants (e.g. by explaining to them the nuances of a debate, how Kialo works, etc.), as well as casual talk (e.g. who has been on vacation/holiday recently, whether someone has gotten busy at work, and so forth). Claim chats tend to be focused more so on the issues with a particular claim (e.g. whether it is unclear, unsupported, or irrelevant), though people discuss higher-level issues here, too. Both discussion and claim chat records are publicly visible, and we collect and organize them as part of our thick record.

Data Analysis

We iteratively read and discussed our thick record, which drew our attention to the way moderation practices changed when Kialo rolled out design updates. In particular, we became interested in the ways that moderators came into conflict with each other as result of those changes. This led us to examine our thick record through the lens of "conflict" and to consider the ways in which conflict could be said to detract from or contribute to moderation practice. We continued our observation work as we performed data analysis, and became aware of the importance of claim vetting, which, in turn, led us to re-examine our data in terms of how conflict between moderators affects claim vetting.

5. Findings

First, we describe how claim vetting involves argumentation between moderators. Second, we describe how constructive dialogue between moderators can produce higher quality claims. Higher quality claims can mean that the claims are clearer, that they have stronger support, or that they become more relevant to a parent claim or main thesis. Although the interactions

and text we describe are publicly visible on Kialo, we have changed all user names and edited text in an effort to maintain user privacy.

5.1. Claim Vetting Involves Argument Between Moderators

Conflicts can arise when one moderator initiates a discussion about a claim and another moderator accepts it before there has been any resolution to the discussion. Since Kialo does not have an official policy on conflict management, moderators take different approaches in response to what they see as a conflict.

An illustrative case in the climate change debate, for example, played out between several moderators across multiple claim chat threads. It began with what one moderator perceived as a breach of protocol by another. @sodanotpop had been workshopping a suggested claim with an author when another moderator, @blueteam, accepted the claim into the debate. @sodanotpop subsequently flagged the claim and engaged @blueteam: “[It] was not appropriate to accept a suggestion still under discussion. i engaged the author in order to strengthen it before accepting it into the debate.” This comment initiated a lengthy argument that played out in three separate claim chat threads, which meant that these two moderators were moving to different claims in the debate arguing with each other about the proper protocol for collaborative claim vetting.

Some of this argumentation was pertinent to the claims themselves. For example, @blueteam discussed newly provided support as justification for accepting claims. “I accepted it because the author’s claim was cited as unsupported, they then supported the claim so i marked it as supported.” They questioned the grounds for other claims. “Where is the evidence or anything else substantial that backs up this claim?? there isn’t any.” Similarly, @sodanotpop pointed out that “the claim contains a link to scientific work that has been disproven (shown to be false) by other members of the scientific community.”

However, they also argued over how to go about collaborative claim vetting. Whereas @blueteam felt justified in accepting a claim that had been marked and was apparently in the process of being workshopped, @sodanotpop believed that it was inappropriate for another moderator to accept a claim that they were workshopping. @sodanotpop could have been echoing the perspective of another moderator in the debate, @libre, who, in a separate thread, called out a user for accepting a claim under discussion. “[I] think it might have been better to not accept this when @saskatoon @sodanotpop and me discuss it.” This comment did not lead to a long argument between moderators. In fact, the person who accepted @libre’s claim did not respond again in the thread.

This could have been due to the way @libre “softened” their comment by acknowledging Kialo’s interface change. “That’s a fairly recent change Kialo seems to have made and we’re still trying to figure out how to use it best.” Following this comment, @libre’s attention returns to the suggested claim. However, in their thread, @sodanotpop and @blueteam continued to argue about moderation policies, including the use of more blunt and direct criticisms.

For example, @blueteam suggested that @sodanotpop did not have a solid grasp of their actions as a moderator on Kialo: “You’re not comprehending what you’re doing here.” They also trivialized an objection “You then had another gripe with the supporting evidence - for another reason, so mark it again,??? so what?” and characterized several accepted claims as fantasies that were “out of touch with reality.” At some point during the latter stages of their argument, @sodanotpop added the following comment to the discussion chat:

*“To all the mods: we need to discuss accepting suggested claims. There are several such claims where I’ve initiated discussions with the authors in order to fix problems *before* accepting them into the debate only to have another moderator come along and accept the claim. This isn’t a great way to collaborate, nor is it a good way to grow the debate. So, I think we need to agree that best practice is to check and see if another moderator has already started discussing a suggested claim before clicking accept. Assume that there is reason that other moderator didn’t accept the claim and at least engage in some discussion before acting.”*

This proposal parallels an earlier insight that @Choco shared with us in a separate debate when we asked about the process of flagging claims. *“Any editor can unflag a claim, but it [sic] generally accepted as a practice that the one who flags it should unflag it.”* After Kialo changed its interface, we have seen some moderators apply this protocol or some version of it. For example, we have seen moderators propose and vote on changes to suggested claims and flagged claims. However, we have also experienced firsthand and witnessed multiple cases of moderators accepting claims without waiting for resolution to an active discussion.

@sodanotpop’s proposal did not result in any broader discussion amongst the moderators with regard to collaborative claim vetting. However, @blueteam posted what seemed to be an antagonistic general critique of how some moderators in the debate vet and accept claims:

“Claims are being accepted that are bordering on insane, with no thought or logic behind them.. for example saying that ‘technically/practically, the tools [to provide energy without burning fossil fuels] exist... in principle there is no obstacle to [stop] burning fossil fuels’.. i mean really?? come on guys. its like just claiming martians put CO2 in the atmosphere.”

After this comment, @sodanotpop seemed to withdraw from the debate. Though they still appear as moderators, they have not accepted claims or participated in discussions of flagged claims since around the time their argument with @blueteam wound down. It does not appear that any other moderators took up the discussion of “best practices” for collaborative claim vetting. However, it is possible to find many examples of arguments between moderators—especially in the climate change debate—where the arguments seem not to affect claim under discussion.

Instead, they involve rhetoric like *“what a blustering bunch of nonsense,” “it’s completely useless to discuss with you...,” “I’ll explain in case anyone with an open mind is reading...”* and *“all you’ve done is change a definition to suit your narrative. whatever.”* At no point does anyone interject or attempt to mediate this exchange, which, as of this writing, is still active, and the claim from which it stems remains flagged and unedited.

5.2. Constructive Dialogue Between Moderators Produces Higher Quality Claims

Dialogue is the primary way moderators resolve issues pertinent to the overall structure of a debate, to particular (problematic) claims, and to moderation practices. While it is possible, in our experience moderators rarely work in isolation. In fact, the two most important elements of the interface might be the discussion and claim chats since these provide the forums for moderator dialogue. Two important features of these chats are: (1) they are

public and thus visible to the entire Kialo community and (2) they are continuous. Public visibility may strengthen civility between participants on the site, and a living historical record provides insight into how ideas may have evolved over time.

With regard to the process of evaluating and accepting suggested claims, the interaction between participants was neither public nor continuous. Moderators requested revisions to suggested claims via direct messages, which were invisible to the broader community. If an author suggested a new claim, the evaluation process began anew. Assuming the same moderator examined the new claim, they would have to recall (from memory) the previous claim as well as their revision request. There was no accessible living record of this interaction.

We are unsure of when exactly Kialo changed things, since we did not receive any formal communication describing changes to the platform. However, in early 2018 it became possible to interact with suggested claims as though they had already been accepted into the debate. Kialo made it so that moderators could make public comments on claims in a continuous chat thread. Aspiring participants could make revisions based on these requests or they could engage with moderators in a dialogue about why they might not want to make a revision. Moreover, multiple moderators could see and join the evaluation process, which created the conditions for dialogue to emerge around suggested claims.

For the most part, these dialogues are productive exchanges of ideas. Moderators weigh in with their thoughts on a particular claim, ask others for their perspective, and render judgments on suggested claims that can be taken into consideration by others when deciding to accept a claim or continue to workshop it. In the following exchange, for example, @libre solicits perspectives from another moderator (@qed) about a suggested claim from a new user @tennisC:

@libre: @qed what do you think, and @tennisC why do you think the parent is unrelated?

@hollywood: @libre @tennisC Interesting. Either this is a con to the parent claim as that claim says the parent does not address the Thesis... I understand the fact that the Claim [sic] "The earth is fine" is right but does not 'con' the parent. I think its fine to argue that the claim is irrevelent or is not a good Con to the thesis

@hollywood: @canoe Rewrite and explain why the parent isn't good in the context of the thesis and I vote to accept

@sodanotpop: @reply the parent rebuts the thesis if we accept that the ends of fighting climate change ought to be [preservation]... There are underlying issues that could stand to be teased out.

There are many examples of dialogue between moderators and writers resulting in concrete improvements to the clarity, relevance, or grounding of a claim. These dialogues tend to include civil language and a respect for other perspectives and approaches — even those that deviate from site-wide conventions for conduct. In a debate about gender as a social construct, for example, someone changed the form of the main thesis without consulting others who had been actively working in the debate. This resulted in a discussion of the merits of the change and, ultimately, a decision to revert the thesis back to its previous form:

@originator: I'll tag @jolene @abcdefg and @grasshopper to see if they agree with the changes.

@jolene: *Some of the reasons expressed have a point. But, I feel the first formulation was clearer for most readers (with little background knowl) and as objective as possible*

@abcdefg: *I think the current wording communicates that gender and sex are the same, and the suggested claims just now coming in reflect this.*

@abcdefg: *I'm going to re-draft it similar to the original for now; we can continue discussing this to get something stronger. Hope that's okay!*

Kialo currently hosts several debates addressing potentially divisive issues, such as the current “stand or kneel” NFL controversy in the United States, abortion rights, and racial profiling. It is understandable that participants in these debates, including moderators, would have strong feelings about these topics. Furthermore, it is also understandable that these feelings would in some way inform their interactions with others on the site. For writers, this might mean posting more “pros” in support of a topic in accordance with their views. For moderators, this could mean holding certain sides of a debate to higher standards as one debate participant suggested: *“This is a clearly biased discussion. You have multiple pro claims that have no support and most of the skeptical ones are challenged repeatedly (to the point that the average contributor would give up).”* Such bias is perceived as a liability—not a strength—on account of how it excludes certain perspectives from the debate.

While there can be drawbacks to moderators having different points of view, it is not necessary to frame points of view as liabilities. There are examples of how different—even opposing—points of view can be used to strengthen claims and debates on Kialo. On the other hand, there appear to be more scenarios involving clashing points of view that devolve into arguments that lead to no concrete improvements in a debate. In some cases, arguments have concrete, negative consequences: participants may withdraw from a debate or decide to stop using Kialo altogether. A key seems to be managing different points of view to facilitate constructive dialogue between adversarial points of view.

6. Discussion

Conflict is possible any time humans interact with each other. It can become problematic if it takes attention away from the task(s) at hand and negatively impacts users’ experiences in an online community. Our experience as participant observers on Kialo motivated us to explore the possibility that showcasing moderators’ potentially conflicting beliefs and values on Kialo might be conducive to constructive dialogue rather than the kinds of argument we observed between @sodanotpop and @blueteam. Thus, we suggest that making moderators’ perspectives on a debate topic more visible is an effective strategy to facilitate constructive interactions between moderators. In addition, we discuss the value of providing a space for moderators to document and iterate on their processes.

6.1. Foreground Potential Sources of Conflict

One interesting and potentially valuable feature on Kialo is the ‘Perspectives’ tool. The tool aims to enable participants to see debates from other participants’ perspectives, including moderators. Participants can cast a vote on the ‘veracity’ of the main thesis (whether they agree) and the ‘impact’ of the claims in the debate (whether a claim effectively supports or

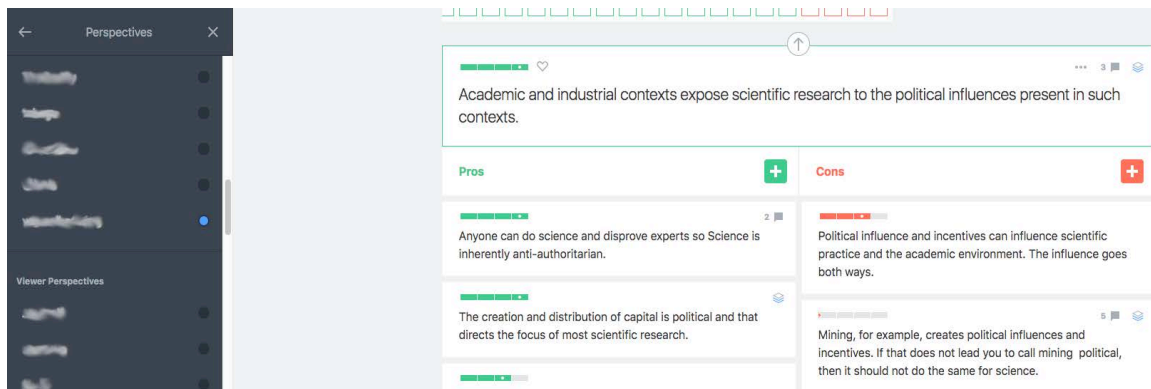


Figure 2: The Perspectives Tool Shows How Users Have Voted on Claims in a Debate.

rebutts the main thesis). The combination of votes on the main thesis and claims forms a participant’s perspective on the debate. We believe that there are several issues pertinent to the perspectives tool.

Voting is not mandatory, and so many participants simply do not vote. Thus, it is not possible to see their perspective. Moreover, it seems as though it is possible to opt out of sharing a perspective. After observing the clash between @sodanotpop and @blueteam, we became interested their perspectives on the debate. We were able to see how @sodanotpop voted on several claims, and the votes would suggest that they were in favor of the main thesis. However, when we tried to see the debate from @blueteam’s perspective, we were unable to locate their user name on the list of active participants. We presume this means there is a way to ‘opt out’ of sharing perspectives. Finally, accessing and making sense of the perspectives tool is not intuitive nor is it efficient. It is located in the discussion menu, which is difficult to find. There is no indication of the tool’s function. Even if someone does filter a debate through another participant’s perspective (Figure 2), they would then have to navigate the debate claim by claim in order to ascertain that perspective. Navigating a debate, even a small one, can be a time-consuming process. If perspectives are important, and we believe they are, then we argue for amplifying them—especially as they pertain to debate moderators.

Knowing other moderators’ points of view can be useful in dialogues about suggested claims. For example, if we know that another moderator strongly disagrees with the claim, “Humans should act to fight climate change,” we could calibrate our comments and questions to take this into account. One reason we grew frustrated with another moderator was that they stated, on multiple occasions, that reason and logic guide their claim vetting process. This gave the impression of a disinterested third party committed to equitable claim-vetting on both sides of the debate. However, as we continued to collaborate with them, we interpreted their actions as favoring one side of the debate and subtly undermining the other. When we raised this publicly, they were quick to rebut our interpretation.

On the other hand, our reluctance to accept their stance could indicate that awareness of other moderators’ points of view might hamper dialogue. Our belief that another moderator acts on the basis of a point of view – as distinct from reason and logic – might undermine our ability to see that, in some cases, their arguments about certain suggested claims are reasonable. We have drawn conclusions about @blueteam’s standpoint, for instance, on the basis of a limited set of interactions that do not represent the whole of their experience.

However, these conclusions influence our interactions. We are more confrontational, which could be good or bad. In some cases, our confrontations have resulted in stronger claims being accepted into the debate. This is good, and it reflects ways in which conflict can be managed and leveraged as a resource. Understanding other moderators' perspectives seems critical here. In others, however, they have fueled circular discussions that do not yield any concrete changes to suggested claims. Instead, they have seemingly become personal and stalled, thus resulting in no changes to the pro or con claims.

Public facing claims about reason and logic as primary motivating factors can minimize the influence that a personal standpoint has on moderation practice. We believe that it is neither possible nor desirable for moderators to operate independently of their perspectives. Instead, we believe that moderators can leverage these perspectives to cultivate better debates and engage in more productive discussions. We are not suggesting that surfacing personal beliefs and values about relevant topics would be a magic bullet for conflict management peer production communities. However, our own experiences on Kialo suggest that surfacing personal interests can be an asset in some cases while in others it might be seen as detrimental.

Kialo could require moderators to vote on the main thesis in a debate to indicate whether they agree or disagree. That stance could then be visualized in a public-facing way so that in any interaction with other moderators or writers it would be possible to glean the moderator's standpoint on the debate topic. Moderators need not be "locked in" to their vote. Other researchers have implemented voting systems before and after participation in a debate (Kriplean et al. 2011), which is a reasonable approach. Moderators could vote on a main thesis as often as they like with new vote informing a change in their standpoint visualization in real-time.

6.2. Toward Constructive Dialogue Between Adversarial Points of View

We have observed and interacted with moderators whose interests seem to diverge from our own. In one notable case, for example, we observed a clash between moderators that came to a head with the acknowledgement that they reached an impasse whereby neither moderator was prepared entertain or accept a point of view other than their own. The clash seemed to result in one moderator withdrawing from participation in the debate, while the other moderator appears active as of the writing of this paper. Moreover, the claim(s) that served as the sites for this clash remained problematic and unresolved until, after several weeks of inattention, other moderators resumed evaluating and workshopping them. This is a good illustration of what (Filippova and Cho 2016) characterized as ideological issues distracting from the task at hand. Problematic claims in the debate remained unresolved because moderators were arguing about their differing beliefs and values, which leads us to suggest that, given the purpose of moderation on Kialo, this dialogue was not constructive.

A simple way to distinguish a constructive dialogue about a suggested claim on Kialo, then, could be its outcome: do any of the moderators involved recognize or appreciate the limitations of their own perspective? Do they acknowledge that other moderators can or should have different perspectives? Finally, and perhaps most crucially, does the dialogue lead to the revision and acceptance (or rejection) of a suggested claim? It seems reasonable to suggest that, in the end, if a dialogue results in some decision about a suggested claim then it could be said to be constructive. Whether moderators' perspectives change could be a secondary concern.

On the other hand, these exchanges can also be rewarding since they could be seen to stimulate stronger arguments. Moderators who do not know (or care) much about a particular topic are unlikely to convince others to reevaluate their own position. Moderators who know or care a lot – regardless of which side they are on – are, in our experience, more likely to ask questions and raises challenges, which, in the long run might be better for the debate. These kinds of moderators force others to “know their stuff,” develop stronger arguments, and remain engaged even when doing so seems counterproductive or frustrating. So, creating an environment conducive to civil, adversarial interactions between moderators might be quite important, and foregrounding their personal standpoints on an issue could contribute to that goal.

Once we drew conclusions, for example, about how other moderators felt about climate change or about “shadow banning” certain kinds of social media accounts, we modified our expectations and rhetorical approach to arguing about suggested and flagged claims. Moreover, we made the decision to engage with them at every opportunity—rather than withdraw—thinking that other participants on the site would benefit from seeing two conflicting points of view engage with each other in order to strengthen a debate and the broader Kialo community. It might be these kinds of interactions that produce the strongest accepted claims, and thus the strongest dialectics, on Kialo.

However, it is also clear that @sodanotpop and @blueteam were engaged in a conflict that seemed to straddle what Filippova & Cho (2016), citing Arazy, Yeo, and Nov (2013), characterize as “process” and “affective” conflict. It was a process conflict in the sense that both moderators argued about how to vet and accept claims as part of a team (Filippova and Cho 2016, p. 707). It was affective conflict in the sense that both moderators took issue with the other’s personal beliefs and values. Each characterized these as “biases” influencing the other’s participation in negative ways.

Throughout their clash, both moderators commented not only on the content of the claims under examination but also on the expected behaviors of moderators when it comes to vetting suggested claims. @sodanotpop, adopting a perspective shared by some other moderators, argued that the moderator who flags a claim, or initiates a workshop with authors and other moderators, should be the one who unflags or decides to accept the claim. While @blueteam did not disagree with this directly, their actions suggest that they see these decisions as open to the moderator team. That is, if a moderator finds a flagged claim and sees a strong rationale for unflagging it, they can remove the flag without consulting others on the team.

We are not making any judgment about which of these two positions is right but, rather, we want to examine ways that they might be put into constructive dialogue with one another. Remaining firmly committed to one or the other position would seem to run counter to the tenets of dialectical reasoning, the purpose of which is to explore two contrasting points of view about an issue in order to produce new knowledge. This could mean appreciating the richness and complexity of an issue that previously seemed black and white or it could mean changing one’s mind entirely, and this would seem to be one of the underlying purposes of Kialo.

A key issue with regard to process conflict on Kialo is a dearth of policy describing practical strategies for interacting with other users. One thing we took away from observing how moderators negotiate their practice is that these negotiations happen in bits and pieces at different sites throughout a debate. So, there is a great deal of informal policy-proposing

by community members. @sodanotpop and @blueteam, for example, spread their argument between three separate claim chat threads as well as the main discussion chat thread. They made several proposals with regard to what the norms for collaborative claim vetting should be, but they did not compare or reconcile these suggestions. Nor did they compile these suggestions in a location that would be widely visible to other moderators.

The discussion chat, which could in principle serve as such a location, is challenging to parse. Moderators and writers post greetings, introductions, links to relevant external content (e.g. journal articles, YouTube videos, etc.), comments on the debate, and personal details. It is organized in reverse chronology, which is typical of many chat logs. There is no low-cost method for searching or filtering content (e.g. by hashtag). The cost of finding meta-level comments and questions by sifting through all other comments might discourage people from trying. We have experienced this frustration firsthand. But perhaps it would be possible to introduce hashtags as a first step to make the discussion chat searchable by content type or to create a singular meta-discussion for the Kialo community where the purpose is to establish and revise a set of clearly defined “rules” for moderating and writing claims. This might have the dual benefit of mitigating disagreements about how to moderate and, thus, within the boundaries of a debate, keeping dialogue between moderators focused on the quality of theses and claims.

6.3. Future Work

We see value in categorizing different kinds of conflict on Kialo and determining which of these kinds could be beneficial. This would require distinguishing good claims from bad ones, and Kialo already has a framework for this purpose. For example, bad claims are those deemed to be unsupported, unclear, vulgar/abusive, unrelated, not a claim, or duplicative. However, we believe it would be possible and useful to apply a framework such as Toulmin’s (Toulmin 2003; Erduran et al. 2004, pp. 92-97) to assess claim quality, which would be a crucial step in a project exploring the relationship between conflict and claim quality. We are also interested in possible ways to iterate on Kialo’s interface and interaction design to help writers craft stronger, more effective claims. Finally, we have already observed how some moderators characterize their actions as driven primarily by reason and logic. This struck us as an interesting discursive strategy that could be part of a broader project aimed at maintaining the authority and power of their position as a moderator. Moreover, it would be interesting to compare how moderators talk about their process with how they carry it out. Hence, we see value in using discourse analysis to examine moderators’ talk about their role and then to compare this talk with an analysis of their actions (accept/reject decisions, edit decisions, and arguments with other moderators).

7. Conclusion

Kialo is a novel online debate platform supporting teams of moderators and writers in the construction of pro/con debates about different topics of interest. We situate our ongoing participant observation in relation to existing research on moderation and online conflict and dispute resolution. We found that moderators with different points of view can clash over suggested claims or accepted claims that have been flagged as problematic, and we explained that these clashes can result in no ostensible improvements to the debate or to the Kialo com-

munity. In some cases, these clashes can discourage participation in certain debates and even result in participants leaving the site altogether. On the other hand, constructive dialogue between moderators has the dual benefit of encouraging participation and strengthening the quality of debates.

During our study, we observed how an interface change facilitated conflicts between moderators and speculated that one possible source of these conflicts could be a lack of awareness of other moderators' points of view. Moderators themselves frame their approach as driven by logic and reason rather than by beliefs and values even when those beliefs and values seem to become visible through interactions with other moderators. We propose that foregrounding moderators' beliefs and values—by foregrounding their position on a given debate topic—could be an effective way to anticipate and preempt conflict. An important next step in our research will examine the way that moderators use language to construct the role that values and assumptions might play in their actions on Kialo.

References

- Arazy, O., Ortega, F., Nov, O., Yeo, L., and Balila, A. (2015). Functional Roles and Career Paths in Wikipedia. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*, pages 1092–1105, New York, New York, USA. ACM Press.
- Arazy, O., Yeo, L., and Nov, O. (2013). Stay on the Wikipedia task: When task-related disagreements slip into personal and procedural conflicts. *Journal of the American Society for Information Science and Technology*, 64(8):1634–1648.
- Billings, M. and Watts, L. A. (2010). Understanding dispute resolution online. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*, page 1447, New York, New York, USA. ACM Press.
- Boellstorff, T., Nardi, B., Pearce, C., and Taylor, T. (2012). *Ethnography and virtual worlds: A handbook of method*. Princeton University Press, Princeton.
- Burbank, N., Dutta, D., Goel, A., Lee, D., Marschner, E., and Shivakumar, N. (2011). Widescope - A social platform for serious conversations on the Web.
- Carspecken, P. F. (1996). *Critical ethnography in educational research: A theoretical and practical guide*. Routledge, New York.
- Dow, S., Kulkarni, A., Klemmer, S., and Hartmann, B. (2012). Shepherd the crowd yields better work. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work - CSCW '12*, pages 1013–1022, New York, New York, USA. ACM Press.
- Erduran, S., Simon, S., and Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's Argument Pattern for studying science discourse. *Science Education*, 88(6):915–933.
- Filippova, A. and Cho, H. (2015). Mudslinging and Manners. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*, pages 1393–1403, New York, New York, USA. ACM Press.

- Filippova, A. and Cho, H. (2016). The Effects and Antecedents of Conflict in Free and Open Source Software Development. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW '16*, pages 703–714, New York, New York, USA. ACM Press.
- Fréard, D., Denis, A., Détienne, F., Baker, M., Quignard, M., and Barcellini, F. (2010). The role of argumentation in online epistemic communities. In *Proceedings of the 28th Annual European Conference on Cognitive Ergonomics - ECCE '10*, page 91, New York, New York, USA. ACM Press.
- Grevet, C., Terveen, L. G., and Gilbert, E. (2014). Managing political differences in social media. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing - CSCW '14*, pages 1400–1408, New York, New York, USA. ACM Press.
- Huang, W., Lu, T., Zhu, H., Li, G., and Gu, N. (2016). Effectiveness of Conflict Management Strategies in Peer Review Process of Online Collaboration Projects. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW '16*, pages 715–726, New York, New York, USA. ACM Press.
- Keegan, B. and Gergle, D. (2010). Egalitarians at the gate. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work - CSCW '10*, page 131, New York, New York, USA. ACM Press.
- Kim, J., Cheng, J., and Bernstein, M. S. (2014). Ensemble. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing - CSCW '14*, pages 745–755, New York, New York, USA. ACM Press.
- Kittur, A., Pendleton, B., and Kraut, R. E. (2009). Herding the cats. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration - WikiSym '09*, page 1, New York, New York, USA. ACM Press.
- Krieger, M., Stark, E. M., and Klemmer, S. R. (2009). Coordinating tasks on the commons. In *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*, page 1485, New York, New York, USA. ACM Press.
- Kriplean, T., Bonnar, C., Borning, A., Kinney, B., and Gill, B. (2014). Integrating On-demand Fact-checking with Public Dialogue. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing, CSCW '14*, pages 1188–1199, New York, NY, USA. ACM.
- Kriplean, T., Morgan, J., Freelon, D., Borning, A., and Bennett, L. (2012). Supporting reflective public thought with considerit. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work - CSCW '12*, page 265, New York, New York, USA. ACM Press.
- Kriplean, T., Morgan, J. T., Freelon, D., Borning, A., and Bennett, L. (2011). ConsiderIt. In *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems - CHI EA '11*, page 1831, New York, New York, USA. ACM Press.
- Liu, J. and Ram, S. (2011). Who does what. *ACM Transactions on Management Information Systems*, 2(2):1–23.

- Luther, K., Fiesler, C., and Bruckman, A. (2013). Redistributing leadership in online creative collaboration. In *Proceedings of the 2013 conference on Computer supported cooperative work - CSCW '13*, page 1007, New York, New York, USA. ACM Press.
- Park, D., Sachar, S., Diakopoulos, N., and Elmqvist, N. (2016). Supporting Comment Moderators in Identifying High Quality Online News Comments. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, pages 1114–1125, New York, New York, USA. ACM Press.
- Schneider, J., Samp, K., Passant, A., and Decker, S. (2013). Arguments about deletion. In *Proceedings of the 2013 conference on Computer supported cooperative work - CSCW '13*, page 1069, New York, New York, USA. ACM Press.
- Settles, B. and Dow, S. (2013). Let's get together: The formation and success of online creative collaborations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 2009–2018, New York, NY, USA. ACM.
- Toulmin, S. E. (2003). *The Uses of Argument*. Cambridge University Press, Cambridge, updated edition.
- Wenger, E. (1998). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press, Cambridge.
- Zhu, H., Kraut, R., and Kittur, A. (2012). Effectiveness of shared leadership in online communities. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work - CSCW '12*, page 407, New York, New York, USA. ACM Press.
- Zhu, H., Kraut, R. E., Wang, Y.-C., and Kittur, A. (2011). Identifying shared leadership in Wikipedia. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*, page 3431, New York, New York, USA. ACM Press.

Affiliation:

Jordan Beck
 The Pennsylvania State University
 University Park, PA
 E-mail: jeb560@ist.psu.edu
 URL: <https://designinquiry.me>

SocArXiv Website
 SocArXiv Preprints

Preprint
[10.31219/osf.io/cdfq7](https://doi.org/10.31219/osf.io/cdfq7)

<https://socopen.org/>
<https://osf.io/preprints/socarxiv>

Submitted: November 2, 2018
Accepted: November 2, 2018
