

Understanding and Augmenting Debates on Reddit

by

Jeremy Zhengqi Huang

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Information
in the University of Michigan
2023

Thesis Committee:

Professor Eytan Adar
Professor Eric Gilbert

Jeremy Zhengqi Huang

zjhuang@umich.edu

© Jeremy Zhengqi Huang 2023

Acknowledgments

I thank Dr. Eytan Adar for his support while completing the thesis and Dr. Eric Gilbert for serving as the thesis committee member.

Table of Contents

Chapter 1 Table of Contents

<i>Acknowledgments</i>	<i>iii</i>
<i>List of Tables</i>	<i>vii</i>
<i>List of Figures</i>	<i>viii</i>
<i>Abstract</i>	<i>ix</i>
<i>Chapter 2 Introduction</i>	<i>10</i>
<i>Chapter 3 Related Work</i>	<i>13</i>
3.1 Characteristics of Online Debates	13
3.2 Online Discourse Cognition and Behaviors	14
3.3 Designs for Productive Discourse	15
<i>Chapter 4 The Reddit Platform</i>	<i>17</i>
4.1 Discussion Threads.....	17
4.2 Voting System and “Karma”	18
4.3 Sorting Options.....	18
4.4 Flairs.....	18
<i>Chapter 5 Method</i>	<i>19</i>
5.1 Interviews.....	19

5.2 Data Analysis	19
<i>Chapter 6 Findings.....</i>	<i>21</i>
6.1 Debate Culture.....	21
6.1.1 Good Faith vs. Bad Faith Debates	22
6.1.2 Misinformation Motivates Participation in Debates.....	23
6.1.3 Heated Exchanges Lead to Harmful Behaviors.....	23
6.1.4 Clear Reasonings Lead to Critical but Productive Debates.....	24
6.1.5 Anonymity and Transient User Relationships Foster Low-stake Debate Environments	24
6.1.6 Openness to Opposing Views and Self-Fulfilling Behaviors	25
6.2 User Interface and Perceived Affordances	26
6.2.1 Research Before Engagement.....	26
6.2.2 Threaded Format Creates Structure and Formality but with Limitations.....	27
6.2.3 Sorting Options Offered Diverse Ways to Understand and Influence Debates.....	28
6.3 Participants' Design Suggestions for Facilitating Debates.....	29
6.3.1 Browsing.....	30
6.3.2 Engaging.....	33
6.3.3 Socializing	35
<i>Chapter 7 Discussion.....</i>	<i>37</i>
7.1 State of Reddit Debates.....	37
7.2 Designing for Productive Debates on Reddit.....	38
7.2.1 Reducing Cognitive Load with Adaptive Viewing Interface	38
7.2.2 Encouraging Introspection and Critical Engagements	39
7.2.3 Leveraging Community Efforts to Foster Healthy Debate Environments	40
7.2.4 Presenting Signifiers for Norms	41
<i>Chapter 8 Conclusion.....</i>	<i>43</i>

Chapter 9 References..... 44

List of Tables

Table 1. List of participants with the self-reported active subreddits and frequency of engaging in debates.	21
Table 2. Aggregated and summarized design ideas from participants for facilitating Reddit debates.	30

List of Figures

Figure 1. Reddit’s Discussion UI. (A) Original Post (OP) with a text box for entering comments as top-level comments. (B) Top-Level Comments with nested (“child”) comments. .17

Figure 2. Twitter’s Community Note feature. This feature allows certain readers to annotate posts by adding background information to the topic.41

Abstract

Due to its low threshold for participation and user-driven nature, Reddit is one of the most popular platforms for democratic debates. Therefore, understanding these large-scale and dynamic deliberations require thoughtful investigations. We interviewed twelve participants to understand the debate culture, their interactions with the Reddit interface, perceived affordances, and design suggestions for improving the debate experience. Our findings revealed some characteristics of Reddit debate culture, including the impact of good and bad faith deliberations, misinformation motivating participation, and anonymity leading to low-stake debate environments. We also found that heated exchanges might lead to harmful behaviors due to the “disinhibited” online environment, and providing clear reasonings for one’s arguments led to critical but civil debates. Users reflected on their encounters with cognitive biases while engaging in debates on Reddit. Based on the findings, we present design implications for enabling civil and constructive debates online, including reducing cognitive load with an adaptive interface, encouraging critical engagement and introspections, empowering the community to moderate content, and introducing signifiers for positive social norms.

Chapter 2 Introduction

Online platforms have become a popular space for public discourse. Among those platforms, Reddit has emerged as an online platform that facilitates discussions on various subjects, including debates on controversial issues. Reddit offers great potential for fostering democratic debates due to its low threshold for participation, diverse user base, and extensive range of topics. Indeed, Reddit users have initiated numerous grass-root communities dedicated to debates (e.g., *Change My View* [43]). Most of the debates on Reddit are user-driven with low participation thresholds. However, its potential for dynamic and large-scale deliberation comes with great complexity. As an anonymous social media platform, Reddit is more likely to house antisocial and harmful behaviors [6,10]. Moreover, the sheer volume of discussion content poses a challenge for users to browse and digest [41]. Understanding the dynamics of Reddit can help us identify design challenges and opportunities to foster a civil and democratic deliberation platform. However, to our knowledge, there is little research dedicated to understanding these complexities.

In addition, there has been growing interest in augmenting online discussions and deliberations. Prior work has explored using visualization and summarization approaches to organize textual information [17,32,40,41] and designing dedicated deliberation platforms that structure the debate process [19,44] and support exposure to diverse views [30]. However, few studies investigated the designs tailored to deliberations and debates on Reddit. As a popular discourse platform with significant offline implications, it is crucial to examine the applications of possible strategies, interventions, or tools that can cultivate a constructive and civil

environment for debates and deliberations on Reddit. Based on the motivations, we asked three research questions:

- **RQ1:** What are users' perceptions of the debate culture on Reddit?
- **RQ2:** How do users browse, navigate, and participate in Reddit debates? What are the interface elements or affordances relevant to the debate experience?
- **RQ3:** What designs can enhance the debate experience and foster a civil and constructive debate environment?

In the present study, we address the above gaps by conducting interviews with twelve participants to understand the debate culture and investigate Reddit's interface elements and affordances related to debates. Our interviews revealed numerous characteristics of Reddit debates. First, the user encounters with good and bad faith debaters significantly influenced the users' perceptions of debate quality. Second, while frustrating to users, false claims motivated engagements. Users reflected on both positive and negative experiences while debating on Reddit. For example, while users understand that Reddit debates can get emotional and heated, these tense exchanges often evolve into harmful behaviors, significantly hindering their ability to debate constructively. On the other hand, users appreciated debates where debaters offered easy-to-understand and robust evidence even if they held different views. This suggests that some design elements of formal debates can be applied to more democratic, online-based deliberation platforms. Users also reported the presence of confirmation biases, echo chambering, and trench warfare dynamics, where being exposed to opposite views reinforced debaters' own beliefs. Based on participants' reflections, design suggestions, and prior work, we came up with design implications for a civil, constructive, and democratic online debate platform, including (1) adaptive content viewing interface for reducing cognitive load, (2) encouraging critical

engagements and introspective behaviors like fact-checking and self-examining arguments, (3) leveraging communities to highlight high-quality contents and combat malicious behaviors, and (4) introducing signifiers for positive norms.

Chapter 3 Related Work

We present the background and related work regarding debate culture and tools for large-scale discussions and deliberations.

3.1 Characteristics of Online Debates

Prior work has explored the characteristics of online discussions, including debates. Online debates differ from in-person ones in that they lack the behavioral cues that only occur during face-to-face interactions [25], making people less likely to reach a resolution or change their views [23]. Similarly, researchers also pointed out that both *echo chambering*, where individuals only interact with opinions that reflect their own, and *trench warfare dynamics*, where opinions are reinforced through both confirmative and contradictory information, characterized online debates [18]. While The anonymized and asynchronized nature of online debates alleviates the fear of online participation [1], it can lead to less controlled (“disinhibited”) interactions with other users [35], potentially causing antisocial and toxic behaviors like trolling [10] and harassment [6]. Other than the design factors, cognitive (e.g., mood and social anxiety) and contextual factors (e.g., exposure to trolls) can also contribute to harmful behaviors online [10,21,26]. Moreover, when the quality of online debates decreases (e.g., people become uncivil [13]), users’ attitudes tend to be more polarized. Interestingly, the qualities of online debates depend on the platforms and deliberation goals [16,17]. For example, more formal, decision-oriented deliberation platforms, including online consultations, house more productive discussions than highly accessible social media platforms like Reddit [15]. For certain communities (e.g., ChangeMyView), the debate dynamic can be competitive since the goal is to persuade other users. Linguistic choices like using calmer words and listing concrete examples

and social factors like the number of unique counter-arguments can motivate view changes [37]. We extend the study of these dynamics by generalizing the population to all Reddit communities and probing their patterns of engaging in debates.

3.2 Online Discourse Cognition and Behaviors

Online debates encompass various aspects of cognitive and social psychology. Due to the sheer amount of online debate content, navigating and participating in such debates can lead to a high cognitive load. When encountering overloads, users tend to resort to reasoning processes that are quicker but heavily influenced by emotions (“affect heuristic”) [33]. A productive debate requires analytical and logical arguments, which can be difficult to generate in an online environment where user behaviors are less inhibited [35].

Confirmation bias is when people seek information that reaffirms their beliefs and dismiss information that contradicts theirs [27]. This phenomenon can be especially pervasive in online environments as users tend to view content that supports their beliefs and interact with users who share the same views, causing polarization and increasing hostilities [4,36]. However, as mentioned in the previous section, prior work found that users can solidify their beliefs through interacting with others who hold both affirming and contradicting opinions (*i.e.*, trench warfare dynamics) [18]. However, this reinforcement effect is likely to be short-term because exposure to opposite views can eventually be a learning process for users (*e.g.*, a slight modification of ideas) [18]. This learning process can be amplified by exposure to “two-sided” arguments or mildly contradicting arguments [18]. For our study, we investigated the above psychological effects in action by understanding their interactions with other users on Reddit and how these interactions shape their experience in navigating and participating in debates.

Social norms refer to a set of beliefs and attitudes that are shared by a group of people and govern behaviors in groups [1]. Interestingly, people generally fail to recognize the effect of social norms on their behaviors despite their influential role [11]. Social norms can be classified as *injunctive norms*, the perceived “ideal” or expected behaviors in social situations, or *descriptive norms*, the observed, real-world behaviors of other people [12]. Reddit users may choose to engage in an ongoing debate politely because (1) they *think* that other users will approve of them if they make polite comments (injunctive norm) or (2) other people are also talking politely to each other (descriptive norm). Another example of descriptive norms in action is the explicit, written rules of some subreddits like Change My View.

3.3 Designs for Productive Discourse

Prior HCI work proposed various design approaches to facilitate browsing and navigating public discourse. These designs influence the dynamics of democratic discourse. For example, designs that implicate the norm of making thoughtful comments can pivot users towards more critical and thoughtful engagements, even without the presence of such behavior from other users [34]. Another design approach is summarization through crowdsourced [40,41] or automatic approach [38]. The crowdsourced approach assigned users to generate summarizations directly or annotate discussions with labels; the automatic approach primarily used ML techniques to highlight key phrases or statements (extractive) or generate a condensed version of the original statement (abstractive). Helping users browse discussion contents through summaries could reduce cognitive load [41], but its effectiveness in countering confirmation bias was limited [28]. Similarly, visualizations like tree-based graphs could effectively structure the discussion flows to facilitate a more “purposeful” browsing experience [24].

Some researchers implemented their designs into novel, fully functional platforms and examined their effectiveness in facilitating public discourse. ConsiderIt [19], for example, implemented a pro/con technique that encouraged users to reflect on other people's points, assess tradeoffs, and contribute "new points." Its deployment demonstrated the ability to foster positive norms and potential for long-term utility but highlighted issues regarding user trust. Another platform, Poli [30], was guided by user needs regarding political engagement online, such as exposure to diverse political perspectives and sources, information filtering, and self-evaluation of their voices' impact. While novel designs of discussion platforms are well-studied, there are still questions left unanswered. For example, what is the current user experience of debating on Reddit? Equally important, what are Reddit users' design needs and preferences for engaging in productive debates? Compared to platforms like ConsiderIt [19] and Kialo [44], Reddit is more conversational and less structured but significantly more prominent in scale and covers a broader range of topics. These characteristics pose a unique challenge to understanding the dynamics of debates and generating designs that make the discourse more productive and civil.

Chapter 4 The Reddit Platform

Since its foundation in 2005, Reddit has become one of the most popular social media platforms. Reddit provides a space for people to engage in discussions, acquire and share information, and participate in debates on any topic. We summarize several UI features that are related to our study.

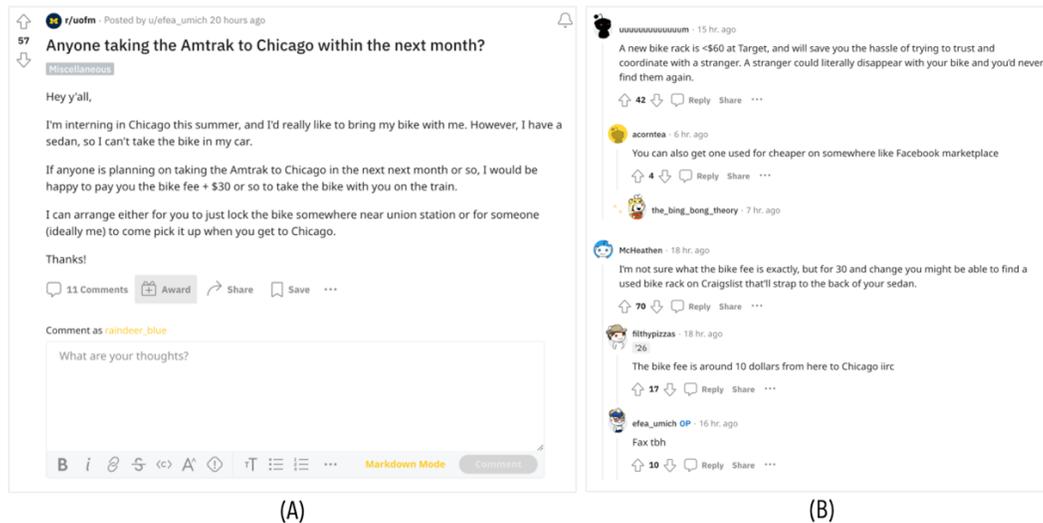


Figure 1. Reddit's Discussion UI. (A) Original Post (OP) with a text box for entering comments as top-level comments. (B) Top-Level Comments with nested ("child") comments.

4.1 Discussion Threads

Reddit organizes and displays all the discussions in a threaded format, allowing users to interact and engage in conversations. The discussion begins with Original Post (OP), a statement or question containing links and pictures. OP determines the context and the topic of the entire discussion. Users can respond to OP with Top-Level Comments (TLC), the direct replies to OP. Nested comments, or replies, are responses to the top-level comments. These follow-ups create a branching structure within the discussions through indenting to visually distinguish from TLCs or other levels of replies. The discussion UI allows users to *collapse* or expand comment threads.

The UI hides all the “child” comments by collapsing a comment. Conversely, expanding a comment shows all the child comments.

4.2 Voting System and “Karma”

Reddit’s voting system plays a significant role in the discussions. Users can upvote or downvote both posts and comments. When the posts or comments get upvoted, the users will gain “Karma” [42]. This scoring system represents the contributions users make to the Reddit communities. Voting can also influence the order of the comments displayed in the thread.

4.3 Sorting Options

Reddit offers various sorting options for the threaded discussions, including sorting by “Best,” “New,” “Top,” “Controversial,” and “Q&A.” Reddit does not officially disclose the algorithms of these options. These sorting options allow a certain degree of customizability for the Reddit viewing experience.

4.4 Flairs

Reddit’s flair system allows users and moderators to tag and classify posts or comments with a descriptive label. Flairs can be assigned to users to indicate roles or to the posts to categorize the content. For instance, a food-related subreddit can assign post tags like “recipe,” “restaurant review,” and “diets.” Each subreddit can customize its flair settings. Therefore, the use and meanings of flairs can vary significantly and depend on the community’s culture.

Chapter 5 Method

5.1 Interviews

Our study was approved by the Institutional Review Board of the University of Michigan. We first posted the study ads in r/SampleSize, a subreddit dedicated to posting surveys and polls, and r/ChangeMyView, a debate-oriented community where users post their views and expect others to challenge their opinions with counter-arguments. We prompted users interested in participating in our study to fill out a survey, which contained questions about their experience debating with other users on Reddit and browsing debate content on Reddit. We then invited 12 participants to semi-structured interviews in Spring 2023. Our participants include users across numerous subreddits and interest groups. We offered \$20 incentives for participation.

The duration of the interviews ranged from 25 to 45 minutes. During the interviews, we asked participants about their detailed experience having debates on Reddit, including (1) goals and motivations for participating in debates, (2) interactions with other users, (3) interactions with the Reddit interface, (4) opinions on affordances and features related to debates, 5) suggestions to improve debates on Reddit. We conducted the interviews through audio or video conference calls on Zoom.

5.2 Data Analysis

We curated our data with participants' survey responses and transcriptions from the interviews. We transcribed our data using Zoom's captioning features. We then conducted an iterative, inductive coding for the transcripts [citation]. We then generated initial codes of the transcripts by summarizing interview notes on a segment-by-segment basis. Finally, we went

through several iterations of the coding process until we could cluster the codes into observable themes.

Chapter 6 Findings

We present the findings from the interviews. Our findings include participants' perceived debate culture on Reddit, their interactions with the Reddit interface, their perceived affordances, and participants' design suggestions for facilitating the debate process.

6.1 Debate Culture

During the interviews, participants shared their experience debating with other users on various Reddit communities (see Table 1). We summarize these experiences to describe users' perceived debate culture on Reddit.

PID	Active subreddits	How often do you have debates?
P1	Politics, World Politics, Political Discussion	Very often (4)
P2	Civil Discourse, ChangeMyView	Often (3)
P3	History, Sports, Algorando Official	Extremely often (5)
P4	Environment, Politics	Extremely often (5)
P5	Debate Religion, Subreddit Drama	Extremely often (5)
P6	Soccer, Politics, League of Legends	Sometimes (2)
P7	ChangeMyView, UIUC	Often (4)
P8	ChangeMyView	Extremely often (5)
P9	Cryptocurrency, Debate	Very often (4)
P10	ChangeMyView	Very often (4)
P11	Debate Like An Englishman	Extremely often (5)
P12	ChangeMyView	Extremely often (5)

Table 1. List of participants with the self-reported active subreddits and frequency of engaging in debates.

6.1.1 Good Faith vs. Bad Faith Debates

Many participants stated that they had encountered users who engaged in debates in both good and bad faith. Participants stated that the good-faith debates include “willingness to accept ideas and different perspectives” (P4), acknowledging “we are not god,” and there is a limitation of individual knowledge (P8) and good intents in general (P7). Good-faith debates motivated the participants to engage deeply with the debates and encouraged a mutual learning process (P1, P3, P12). For example, P12, who actively participated in debates on ChangeMyView, mentioned:

“I engaged with someone in a conversation about gun laws. I shared some facts with OP that he was unaware of. These facts were existing legislation, so it was not really a ‘my source is better than your source’ context, and we each walked away from the convo enriched.”

On the other hand, some participants had experienced debates where other users had no intentions to engage in reasoned or productive discussions. Those bad faith encounters significantly lower the qualities of debates and motivations for engaging in debates. One of the most notable cases participants mentioned was trolling, where users intentionally posted provocative and disruptive comments (P7, P11). Trolling behaviors were P7’s primary reason for her “gradually drifting away” from debates on Reddit, and P11 described that facing trolls made him “feel offended” and “feel like a clown.” Another bad-faith behavior was “malicious blocking and deletion” (P8), where users blocked certain people so they could not reply with counter-arguments or flat-out deleted posts or comments. Some participants mentioned that some Reddit users were “absolutely convinced that they are not affected by confirmation bias” and “cannot appreciate reasonings.” (P12)

6.1.2 Misinformation Motivates Participation in Debates

Many participants mentioned the presence of misinformation as one of their primary frustrations during the debate process (P1, P3, P6, P7). Despite this, other users presenting false claims also became one of the primary drivers to participate in the debates. For example, P1 stated that if he found someone *“saying something that is not true,”* he would *“jump into the topic”* and *“point the false claims out with comments.”* *“I sometimes join those exchanges in the hope that I can clear up some misinformation,”* P7 lamented, *“so many people are posting without any important knowledge.”*

6.1.3 Heated Exchanges Lead to Harmful Behaviors

Emotional and heated exchanges assumed a big part of the Reddit debate experience for many participants. Some participants admitted that they might become emotional when they feel like *“losing an argument”* (P1), and realizing this made them *“feel terrible”* (P5). Furthermore, it is not uncommon for participants to encounter heated exchanges. As P7 explained:

“In CMV, you know you are stepping into an arena. When you post things, you are meant to be challenged. A lot of people do not realize how hard the conversations are and how many people are going to be heated up... sometimes you are facing a sparring match, and you did not even realize.”

Unfortunately, those heated exchanges might evolve into insults and personal attacks. For example, P7 had other users *“DM me from various accounts with nasty comments.”* *“... to the point that I sometimes had to turn off the Direct Messaging feature,”* P7 added. Interestingly, while acknowledging that some topics *“get people easily riled up,”* P11 said, *“I enjoy staying*

calm and logical in responses while they are getting increasingly infuriated.” P11 “considered it a win if my opponent gave up in frustration, even more so if they chose to block.”

6.1.4 Clear Reasonings Lead to Critical but Productive Debates

While Reddit is not designed for formal debates, many participants voiced their desire for arguments where people “put forward their thought process” (P3) and “provide justifications for their views” (P8). “*Sometimes I had to go more than halfway through the post to get to their actual views,*” P8 added, “*I would much appreciate it if people are like, ‘this is my view, and here’s the reason why,’ like... I don’t need to hear three pages of your personal stories.*”

Interestingly, P8 and P12 mentioned that people “listing out the justifications” could help them recognize the “moral systems” of users they debated with (P8, P12) and establish “common grounds” (P12). “*Even for controversial topics, I try to find out which moral framework or principles they are under... so that I can work with them under that framework, like expanding on their points and presenting some alternatives,*” P12 explained.

6.1.5 Anonymity and Transient User Relationships Foster Low-stake Debate Environments

While Reddit is widely recognized as a “social media,” some participants mentioned that the ephemerality of relationships with other debaters was “*what set Reddit apart from other platforms*” (P7). They thought that the “*topic-based*” nature of Reddit debates fostered an environment where users were able to “*participate with interest*” and “*leave freely with no strings attached*” (P6). P10 echoed this: “*I see no reason to make the platform any more ‘social’... honestly, this is what makes Reddit special to me. People come to debate and go.*”

Participants also appreciated the fact that Reddit is an anonymous platform, as user identities are “masked” (P12) by avatars. Such anonymity removed the power dynamics within the

communities and fostered an environment where the quality of individual arguments is more valuable than “user statuses” (P8). P12 stated that this characteristic distinguished itself from other social media sites, explaining:

“You can be Elon [Musk] and someone and post something like ‘go put all the money in Dogecoin’ ... this might work on Twitter, but on Reddit... yeah, good luck with that. People are just going to think that you are a lunatic.”

6.1.6 Openness to Opposing Views and Self-Fulfilling Behaviors

Overall, in terms of perceived openness for constructive debates, participants have reported mixed experiences. Some participants (P8, P12) explicitly stated that they had had a “great experience” debating with other people on Reddit because both were willing to “address limitations of our views” (P12). On the other hand, five participants, including P8, mentioned that they had encountered users who displayed echo chambering behaviors or were not open to having their views challenged. For example, P8, who was an active debater on r/ChangeMyView, stated: *“I have seen beautiful back-and-forth conversations, but I have also seen [Users on CMV] might say ‘Oh, this is my view, and I don’t think any of you will change it... and I will think, well, then why are you even here? This completely defeats the purpose.”* Moreover, P10, who also debated in ChangeMyView, described a more implicit way of avoiding challenging views:”

“One subtle thing I noticed is that OPs usually post their views with certain expectations of where that conversation will head towards. If they saw some unexpected [views], they would just ignore it, which is quite a shame because they were usually some really good ones.”

These experiences reflected the *trench warfare* dynamics [18] in the online environment in that people seeking debates might tend to reinforce their beliefs by either preemptively claiming the superiority of their views or posting their beliefs with expectations of what the opposing views might be. As described by P10, debaters may fall into a self-fulfilling situation where they present a view with a prepared “playbook” of counter-arguments for opposing views and intentionally avoid the opposing views that were not in their playbooks.

6.2 User Interface and Perceived Affordances

We present findings regarding participants’ experience browsing and engaging with debates on Reddit and their evaluations of the user interface and perceived affordances of the Reddit platform related to the debate experience.

6.2.1 Research Before Engagement

Many participants stated that they tended to research the topics before they engaged in the debates. The subjects of research included the overview of the topics (P1, P6), terms and definitions (P3, P6), and the details of certain events (P9). For example, P6 stated: *“If I don’t have any idea what the person was talking about, I do my own research, so I don’t end up spitting out wrong information and being like a fool.”* P1 echoed, saying that *“I will always try to figure out if I have enough information at hand before I go through the comments. If not, I will Google it.”* In addition, some participants held similar standards for other debaters, saying that they would lose interest in engaging with the person they were debating with if they *“turned out to not know what they are saying”* (P8) or *“acting confused themselves.”* (P12)

When asked about the channels for acquiring information, participants reported that they used search engines (P1, P6) or links in posts and comments (P3). P3 noted that the links were

always “external,” so users would need to “*literally leave Reddit*” and “*read through the site*” to find relevant information. This indicates the need for a reliable channel to acquire background information about the discussed topics without disrupting the current flow of engaging in debates.

6.2.2 Threaded Format Creates Structure and Formality but with Limitations

Reddit organizes the discussions on the platform in a threaded format, where each conversation is structured in a tree-like hierarchy of comments and follow-ups or replies. The original post, which can be a link, image, or text, sits at the top of the thread as a “parent” comment. Users can then post comments that become “child” comments and are nested beneath the “parent” comment. Some participants approved of the “thread” format, describing it as “easy to follow” (P1), “well-arranged,” and “useful when having serious debates” (P4). On the other hand, some participants expressed some frustrations with the format, mostly regarding the long comment chains. For example, P6 described the browsing experience as sometimes “stressful” from “getting tired of looking at the rest of replies.” P12 concurred:

“I think it’s 80% good from an information architecture perspective, but where it gets to be a problem... is that when there are so many layers... to the point that it’s basically unusable.”

This indicates that the cognitive load for digesting information in a threaded structure can increase significantly once the volume of information grows. Interestingly, while “irritated” by Reddit’s default behavior of collapsing long comment threads, P7 acknowledged that exposing all the comments in the thread helped “*reveal more about people*”:

“Someone made a post about something involving linguistics and racism. All the surface-level debates were just people arguing about racism... until the debate went about five replies deep, and you realized that the person actually didn’t have any formal education in linguistics and realized that they were completely misunderstanding something.”

6.2.3 Sorting Options Offered Diverse Ways to Understand and Influence Debates

Reddit has several built-in sorting options to enable users to view posts, such as Top, New, and Controversial. Some participants thought that the default sorting (Best) was sufficient for viewing “highlights” of the debates (P4), but others tended to experiment with different sorting options (P7, P10, P12). For example, P12 explained:

“I switch up these options sometimes. They were defaulted to best... so I will read that 10%, and then I will do a sort by Q&A, which is like triggered the most conversations... then I will see what the latest is, because the latest is sort of representative of like... how does the debate evolve in time.”

Similarly, many other participants mentioned that they switched to “Sort by New” because they wanted to see “*the latest development of the debates*” (P7), and newer comments tended to “*not get washed by votes other peoples’ follow-ups*” (P10). “*When I engage with new comments or posts, it is like hanging out with a bunch of guys,*” P12 added, indicating that displaying debate contents in a temporal manner can foster a sense of real-time involvement despite the platform’s asynchronous nature.

However, Reddit did not officially disclose the algorithms of different sorting options, which caused some confusion among users: “*I honestly don’t know what the ‘Best’ is for, though I think they used to explain it.*” (P12)

6.3 Participants' Design Suggestions for Facilitating Debates

During the interview, we asked participants to brainstorm design ideas and suggestions for Reddit to enhance their debate experience. We aggregated those design suggestions and clustered them according to three interaction types on Reddit: browsing, engaging, and socializing (Table 1). Some of the design ideas reflected our findings about the perceived debate culture on Reddit and participants' interactions with the Reddit interface.

Interaction	Design idea	Descriptions
Browsing	Background knowledge	Providing relevant information for users to gain a better understanding of the debate topics. The examples include details of an event and descriptions of a concept.
	Fact-checking	Leveraging AI or community powers to fact-check statements within the debates and potentially correct false claims.
	On-demand thread expansion	Implementing an “on-demand” collapsing of threads and viewing of nested “child” comments instead of displaying all at once.
	Filtering	Providing options to filter debate content based on users, keywords, or topics.
	Highlighting constructive thoughts	Allowing users or moderators to highlight or mark the comment as “constructive” or “thought-provoking.”

	Summarizations	Using tags or summaries to allow users to quickly grasp the key point of posts and comments
	Show deleted comments	Enabling users to view deleted comments by other users for a certain amount of time.
Engaging	Multimedia references	Enabling users to make references beyond attaching links and towards visual, audio, or video sources.
	Warnings for potential insults	Implementing a built-in warning for the textbox if users include potentially harmful statements.
	Exposure to opposing views	Exposing users to opposing views when drafting their arguments.
	Explicit Structuring	Prompting users to explicitly state the assertion with evidence and reasonings
	Trolling behavior detection	Enabling the platforms to monitor if the drafted comments involved trolling-like behaviors
	Socializing	Time outs
Malicious blocking ban		Detecting malicious blocking behaviors (e.g., blocking OP after making a comment)

Table 2. Aggregated and summarized design ideas from participants for facilitating Reddit debates.

6.3.1 Browsing

Many participants (N=9) suggested designs that aimed to facilitate the browsing experience, including:

Background knowledge. As mentioned above, some participants (P1, P3, P5, P6) wanted to make sure that they had sufficient knowledge of the topic before engaging in the debates. Some of their design suggestions echoed this sentiment. For example, P1 imagined the background information to be “accessible” on the page: *“It will be great if that [background] information is accessible... like something below the OP’s description, like, ‘this post talked about blah blah blah...’”* Similarly, P3 suggested that the users should be able to acquire background information without leaving the thread (e.g., switching to a new tab in the browser). Moreover, P6 suggested an in-thread dictionary where users can easily access the definitions of jargon. Presenting background or contextual knowledge about discussed topics has been implemented in mainstream platforms like YouTube [45]. Access to background knowledge or contexts can keep users informed and realize the knowledge gap, potentially framing debaters towards a “growth mindset” [22].

Fact-checking. Two participants mentioned that Reddit should implement a fact-checking feature through moderating (P3) or AI (P7) efforts. For example, P3 suggested having an annotation system for false information that prompts moderators to review:

“If there is misinformation, we should be able to scratch that piece and provide [viewers] with real facts to counter. And people seeing [the thread] will see that whatever this person said is under investigation and should be taken with a grain of salt.”

On-demand thread expansion. Instead of expanding the threads by default, P6 and P7 would like Reddit to collapse the threads (hide child comments) by default and let users decide when to expand and view followed-up comments. P7 corroborated: *“When you go deep in the threads, there often isn’t any more useful information.”* P6 shared the sentiment, adding that:

“I would be interested in seeing how people respond to OP first, but [Reddit] often just displays all the comments, so I will have to scroll for a couple of seconds to the next direct response. It will be much easier if [Reddit] opens those comments only when I want to.”

Filtering. P7 suggested Reddit allows users to filter keywords and users among debate contents to find certain contents or block out contents that users wish to skip. She explained:

“It can go both ways... Sometimes I want to see if anyone else has brought up certain things... but I really want a way to block out things that I don’t want to see. I got so burnt out seeing and explaining the same thing again and again, and these [conversations] rarely get somewhere.”

While filtering the debate content allows users to engage in a more purposeful browsing experience, it can create echo chambers as users may also use it to avoid exposure to opposing views.

Highlighting constructive arguments. Two participants (P10, P12) suggested Reddit enables users or moderators to highlight arguments that are “thought-provoking” and “sparks more meaningful conversations.” In addition to conventional moderating work (e.g., removing rule-violating contents), P10 wanted the moderators to play a “judging role” and highlight constructive arguments. P12 suggested a similar approach, but instead of involving moderators, he suggested the platform to enable the public to “make a note” of constructive content and use algorithmic techniques to prioritize these contents. This finding echoed a previous finding that generalized this approach to all online platforms that involved “difficult” conversations (e.g., disagreements, conflicts, etc.) [3].

Summarizations. Some participants proposed different ways to reduce the burdens of reading long threads. For example, P3 wished the platform can extend the use of *flairs* and apply them to individual comments: *“I think flairs can be used to [extract] the keywords from the comments so I can kind of know what they are talking about [at a glance].”* On the other hand, P8 wanted a manually or automatically generated “TLDR” for posts and comments, saying:

“Some people actually do their own TL; DRs at the end of a very long post or other stuff, but most of them don’t. Honestly, this should be required if you are writing an essay because no one cares about your life story with your dad here... people come here to make a point.”

As mentioned earlier, summarizations can effectively reduce the cognitive load while browsing the contents [41]. However, when it comes to countering the cognitive biases (e.g., confirmation bias), their utility is questionable [28].

Show deleted comments. Currently, Reddit allows users to delete their comments, but doing so will leave a mark showing the deletion. P7 wanted to push this further by partially or fully showing the deleted comments. *“Hopefully,”* P7 added, *“when people see that their trace in this thread is not going away, they will stop making some stupid arguments and just delete it.”* P7 further explained that the reason for this design is to stop malicious deletion of arguments: *“Some people like to post something provoking and delete it later to make other people look bad.”* While showing deleted comments may stop users from posting thoughtless or provoking arguments, this approach raises the stake of participation and overlooks other potential reasons users delete their posts (e.g., lack of “instant social cues” to assess the norms and impact of the post [5]).

6.3.2 Engaging

Many participants suggested designs that facilitate participation in debates and prevent harmful behaviors before it influences the dynamics of the debates.

Multimedia and easy-to-access references. Currently, Reddit allows users to attach links to external sites or add source information to their comments if they want to provide references for their arguments. However, three participants (P1, P3, P6) stated that the platform should make it easier to point to references or sources that support their arguments. For example, P3 described his design idea:

“Not sure if this will actually work, but what if you can click on some link, and it will be some pop-ups that show the quotes from some website and stuff?”

P6 echoed this idea, saying that debaters should be able to “drop a snippet of [source information]” so the task of looking up information will not interrupt the flow of engaging with debates.

Warnings for harmful behaviors. To address the issue of malicious behaviors like trolling and harmful language, three participants suggested that the platform implement designs that warn or stop users’ postings if they enter arguments or comments that may decrease the quality of the debates. For example, P4 described a design that would “disable the post button” if the system detected that the user wanted to post trolling or “insulting comments.” P6 imagined a more explicit warning in the textbox informing the users about the potential harm of their comments. However, this approach needs to be carefully evaluated as users’ willingness to implement this design lied on the borderline [3].

Exposure to opposing views. Given that participants mentioned the existence of trench warfare and self-fulfilling behaviors in debates, it is not surprising that some participants suggested designs that actively expose users to opposing arguments or views. For example, P12 proposed a design

that assigns a dedicated role, “opposer,” whose main contribution to the debate is to “break the norms” and actively oppose debaters’ views. This design is based on the idea that “*good ideas and truths should be resilient*” and that high-quality arguments should be “*tested against bad ideas.*” The design most similar to this suggestion is *Poli* [30], which intelligently exposes users to diverse information based on their browsing habits and preferences. However, prior work contradicted this suggestion, finding that exposure to opposing political views on social media could cause more polarization [2].

Scaffolding the reasoning process. Motivated by the desire to reduce the cognitive load to digest other people’s arguments, some participants suggested designs that help debaters structure their arguments. For example, P12 proposed a tool that automatically generated a “bullet-point list” of the arguments based on the user inputs and explained: “*I think this is a great way to make people critically examine their beliefs like is my point really making sense?*” P12 suggested a more proactive approach: implementing a template system for organizing debates:

“I posted this idea about ChangeMyView to another subreddit a couple of months ago... for serious debates, they should have a template where people can fill in things like arguments or counter-arguments with [listed] evidence.”

6.3.3 Socializing

Other than browsing and engagement-related designs, some participants proposed social features that aimed to make the debates on Reddit more civil.

Time outs. P4 and P6 proposed “time-outs” for users demonstrating malicious behaviors like “personal attacks” and trolls. For example, P4 explained: “*Some people on Reddit need to be treated like kids. Like, if you acted like an a**, you go stand at the corner for 24 hours.*”

Detecting malicious banning behaviors. Based on her observation that some users tended to “drop an argument and ban the users from replying to it,” P7 suggested Reddit deploys a system that automatically detects such malicious banning from users.

Chapter 7 Discussion

Overall, participants stated that they actively engaged in debates on Reddit to exchange ideas, learn new things, and practice reasoning and debate skills. Through semi-structured user interviews, participants identified several important characteristics of Reddit debate culture and reflected on their interactions with the Reddit interface and other users during debates.

Moreover, many design suggestions proposed by participants echoed their reflections on the Reddit debate experience, ranging from browsing to engaging and socializing. Here, we discuss the state of debates on Reddit and point out the design implications stemming from the study.

7.1 State of Reddit Debates

Overall, participants appreciated Reddit as a platform for debates because of its anonymity, asynchrony, and transience of user relationships that distinguish itself from other social media platforms like Twitter and Facebook. However, the sentiments toward its effectiveness in fostering productive debates are mixed. The qualities of debates could be influenced by human factors like background knowledge, logical and reasoning abilities, and emotions. These factors can be the catalyst for productive interactions with other users. For example, many participants tended to be engaged and invested in the debates if others presented and provided sound justifications for their arguments. However, as participants mentioned, these “fruitful” interactions were relatively rare, given that user behaviors tended to be less “inhibited” in anonymous online spaces. Notably, participants’ perceptions of the debate culture on Reddit culture were swayed by other users’ behaviors, which reinstated the influential role of social norms in online settings [14]. For example, the majority of the participants listed trolls, insults, and bad-faith discussions as the primary reasons they refused to engage in the debates. While

misinformation caused frustrations, it was marked as one of the participants' motivations to get involved (e.g., downvoting and commenting), indicating that debaters on Reddit sought impacts across stages of participation. This finding supports prior work on "hard arguments" online [3] and indicates opportunities to leverage community efforts to combat misinformation.

7.2 Designing for Productive Debates on Reddit

We present the participants' evaluations of current designs and design implications for democratic, social media-based debates based on participants' reflections on their debate experience on Reddit and their design ideas for facilitating productive debates.

7.2.1 Reducing Cognitive Load with Adaptive Viewing Interface

During the interviews, participants reported that they invested different amounts of cognitive effort across stages of engagement. For example, some of them skimmed the threads first and decided whether to dive into the discussions based on their judgments of the debate quality and personal interests. At this stage, participants wanted to quickly extract OP's (original post and posters) views and assess their familiarity with the topic. They also sought to capture the dynamics of the debates and stayed up to date with the latest development quickly. When they started diving into the individual comments within the debates, they had a higher tolerance for a slower but more critical browsing experience. Still, the transient nature of interactions and the high volume of content posed challenges like cognitive overloads, indicating that online debate platforms with large user base should implement a more adaptive approach to organize contents, where the interface show details of arguments "on-demand" and give users more controls of the flow and presentations of content streams. There are numerous ways to implement this approach, including those explored by prior HCI work. For example, forums that use threads to present

discussions can leverage *recursive summarization* [41] to iteratively apply summarizations to different levels of threads and build an expandable “summary tree.” A summary tree, if accurate, can provide viewers with “at-a-glance” information about the debate dynamics and make users’ System 1 thinking process more efficient. Another potential approach is interactive visualizations. This approach grant users the ability to overview, filter, and view nested information (e.g., follow-up comments or arguments) on-demand [31], which allows users to browse more purposefully and discover relationships among the arguments [24]. It is important to maintain the ease of use of the interface while implementing the above approaches since the summative components should not disrupt the reading flow or complicate the UI.

7.2.2 Encouraging Introspection and Critical Engagements

Introspection is a cognitive process in which we examine and learn about our current thinking, beliefs, and reasonings [29]. Across the stages of engaging with Reddit debates, users have demonstrated introspective behaviors like assessing their knowledge about certain topics and reviewing the drafts while expecting other debaters to do the same (e.g., fact-checking). Prompting introspection is not a novel idea: Bossens et al. [8] explored a proactive approach in the comment section of news platforms, where the interface nudged users to provide reasonings for agreeing and disagreeing with the statements, and found that it increased the quality of discussions. Moreover, L. de Carvalho et al. [20] explored prompting users to self-report their emotional states before leaving comments, which led users to self-reflect on their inputs in online space. However, these proactive designs instated extra steps to engage with online discussions, which were likely to be rejected by users seeking fast-paced interactions [20]. Therefore, future designs aiming to encourage critical reflections and introspective behaviors on democratic debate platforms should place minimal cognitive burdens on the users to ensure high adoption and high

impact. Furthermore, designs that prompt a *moderate* amount of reflection are more likely to elicit arguments that leverage the strengths of intuition and analytical thinking [7].

7.2.3 Leveraging Community Efforts to Foster Healthy Debate Environments

Currently, moderators are the primary forces to impose rules and regulations on Reddit debates, though users can report malicious behaviors. Previous research found that heavy moderation can demotivate users from deliberating due to the potential perception of censorship [39]. The moderating effort becomes more complicated when the rules of engagements are not clearly written, leading to them being up to moderators' interpretations and causing imbalances in power dynamics within the communities. During the interviews, the majority of the participants expressed their willingness to contribute to the moderation tasks through low-cost and low-effort annotations, which echoed the overwhelming support Brewer et al. received with a Twitch initiative that leverages community effort to address harassment [9]. This finding suggested opportunities for designs that democratize the moderation process through low-cost and reliable crowd annotations. The annotations can serve to label misinformation and identify harmful behaviors [37], but their capabilities can be expanded to highlight arguments that are insightful and constructive [3]. While a similar approach is implemented on Twitter [46] (“Community Notes”), little has known about its real-world effectiveness in combatting false information. Moreover, it would be interesting to explore how moderators, community members, and AI technologies like GPT [47] can collaboratively identify, investigate, and prevent misinformation.



Figure 2. Twitter’s Community Note feature. This feature allows certain readers to annotate posts by adding background information to the topic.

7.2.4 Presenting Signifiers for Norms

When engaging in Reddit debates, users can form a perception of the normative conduct through observations of the behaviors within the communities (“descriptive norm”) and inferences of the approved or ideal behaviors (“injunctive norm”) [11,12]. While descriptive and injunctive norms are distinct concepts, the two can overlap with each other: if users are exposed to bad-faith debates or malicious behaviors like trolls and insults, their perception of how to approach the debates may be swayed towards the toxic side. While the low-threshold nature of Reddit may struggle to hold off malicious behaviors, the platform, along with community leaders like moderators, can embed signifiers that provide transparency to the consequences of individual actions and proactively nudges users toward civil conduct. Compared to interventions, signifiers are more light-weight and are less likely to distract users from engaging in the debates because it does not require any user actions. Moreover, while interventions can steer user behaviors, they induce user concerns such as the intrusion of privacy and “dilution” of their

intended messages [3]. Therefore, these signifiers should foster an environment that induces assumptions of civil behaviors from Reddit debaters with minimal distraction to the engagement with the threads.

Chapter 8 Conclusion

This work used semi-structured interviews to understand the debate culture on Reddit and investigate the users' experience engaging with debates on Reddit. We also encouraged the users to generate design suggestions based on their reflections on the debate experience. Overall, participants' sentiments toward the debate culture were mixed, as a majority of them have experienced both constructive and "bad faith" arguments (e.g., trolls). We also observed that the behaviors of other users significantly influence the perceived debate quality. For example, the increasing amount of "bad faith" comments and heated exchanges prevented users from productively engaging with debate content, while others' thoughtful arguments encouraged users to be analytical and adopt growth mindsets [15].

In terms of user interactions, we found that users tended to self-research to ensure sufficient knowledge about the discussed topics before engaging in the debates. While users generally approved the threaded organizations of asynchronous discussion content, they reported increased cognitive load for longer debates, indicating the need for more adaptive and interactive interfaces (e.g., on-demand expansion) for displaying debate content. Reddit's sorting options enabled diverse ways to engage with the debates, but the algorithms for these options remain undisclosed.

Participants brainstormed numerous design suggestions for better browsing, engaging, and socializing experiences. We analyzed their proposed designs and reflections on the debate experience and concluded with four design implications for augmenting productive and democratic debates: (1) implementing an adaptive interface to reduce the cognitive load of browsing the debate, (2) encouraging introspection and critical engagement, (3) leveraging community effort to foster healthy debate environments, and (4) introducing signifiers of social norms to encourage constructive arguments and civil behaviors.

Chapter 9 References

- [1] Icek Ajzen. 1991. The theory of planned behavior. *Theor. Cogn. Self-Regul.* 50, 2 (December 1991), 179–211. DOI:[https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- [2] Christopher A. Bail, Lisa P. Argyle, Taylor W. Brown, John P. Bumpus, Haohan Chen, M. B. Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky. 2018. Exposure to opposing views on social media can increase political polarization. *Proc. Natl. Acad. Sci.* 115, 37 (September 2018), 9216–9221. DOI:<https://doi.org/10.1073/pnas.1804840115>
- [3] Amanda Baughan, Justin Petelka, Catherine Jaekyung Yoo, Jack Lo, Shiyue Wang, Amulya Paramasivam, Ashley Zhou, and Alexis Hiniker. 2021. Someone Is Wrong on the Internet: Having Hard Conversations in Online Spaces. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1 (April 2021), 1–22. DOI:<https://doi.org/10.1145/3449230>
- [4] Alessandro Bessi, Fabiana Zollo, Michela Del Vicario, Michelangelo Puliga, Antonio Scala, Guido Caldarelli, Brian Uzzi, and Walter Quattrociocchi. 2016. Users Polarization on Facebook and Youtube. *PLOS ONE* 11, 8 (August 2016), e0159641. DOI:<https://doi.org/10.1371/journal.pone.0159641>
- [5] Parantapa Bhattacharya, Saptarshi Ghosh, and Niloy Ganguly. 2022. Analyzing Regrettable Communications on Twitter: Characterizing Deleted Tweets and Their Authors. Retrieved June 28, 2023 from <http://arxiv.org/abs/2212.12594>
- [6] Lindsay Blackwell, Jill Dimond, Sarita Schoenebeck, and Cliff Lampe. 2017. Classification and Its Consequences for Online Harassment: Design Insights from HeartMob. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW (December 2017), 1–19. DOI:<https://doi.org/10.1145/3134659>

- [7] Lisa Bortolotti. 2011. Does reflection lead to wise choices? *Philos. Explor.* 14, 3 (September 2011), 297–313. DOI:<https://doi.org/10.1080/13869795.2011.594962>
- [8] Emilie Bossens, Elias Storms, and David Geerts. 2021. Improving the Debate: Interface Elements that Enhance Civility and Relevance in Online News Comments. In *Human-Computer Interaction – INTERACT 2021*, Carmelo Ardito, Rosa Lanzilotti, Alessio Malizia, Helen Petrie, Antonio Piccinno, Giuseppe Desolda and Kori Inkpen (eds.). Springer International Publishing, Cham, 433–450. DOI:https://doi.org/10.1007/978-3-030-85610-6_25
- [9] Johanna Brewer, Morgan Romine, and T. L. Taylor. 2020. Inclusion at Scale: Deploying a Community-Driven Moderation Intervention on Twitch. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, ACM, Eindhoven Netherlands, 757–769. DOI:<https://doi.org/10.1145/3357236.3395514>
- [10] Justin Cheng, Michael Bernstein, Cristian Danescu-Niculescu-Mizil, and Jure Leskovec. 2017. Anyone Can Become a Troll: Causes of Trolling Behavior in Online Discussions. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, ACM, Portland Oregon USA, 1217–1230. DOI:<https://doi.org/10.1145/2998181.2998213>
- [11] Robert B. Cialdini. 2005. TARGET ARTICLE: Basic Social Influence Is Underestimated. *Psychol. Inq.* 16, 4 (October 2005), 158–161. DOI:https://doi.org/10.1207/s15327965pli1604_03
- [12] Robert B. Cialdini, Raymond R. Reno, and Carl A. Kallgren. 1990. A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *J.*

- Pers. Soc. Psychol.* 58, 6 (June 1990), 1015–1026. DOI:<https://doi.org/10.1037/0022-3514.58.6.1015>
- [13] Kevin Coe, Kate Kenski, and Stephen A. Rains. 2014. Online and Uncivil? Patterns and Determinants of Incivility in Newspaper Website Comments. *J. Commun.* 64, 4 (August 2014), 658–679. DOI:<https://doi.org/10.1111/jcom.12104>
- [14] Sara De Candia, Gianmarco De Francisci Morales, Corrado Monti, and Francesco Bonchi. 2022. Social Norms on Reddit: A Demographic Analysis. In *14th ACM Web Science Conference 2022*, ACM, Barcelona Spain, 139–147. DOI:<https://doi.org/10.1145/3501247.3531549>
- [15] Carol S. Dweck. 1999. *Self-Theories : Their Role in Motivation, Personality, and Development*. Taylor & Francis Group, London, UNITED KINGDOM. Retrieved from <http://ebookcentral.proquest.com/lib/umichigan/detail.action?docID=1582002>
- [16] Katharina Esau, Dannica Fleuß, and Sarah-Michelle Nienhaus. 2021. Different Arenas, Different Deliberative Quality? Using a Systemic Framework to Evaluate Online Deliberation on Immigration Policy in Germany. *Policy Internet* 13, 1 (March 2021), 86–112. DOI:<https://doi.org/10.1002/poi3.232>
- [17] Katharina Esau, Dennis Friess, and Christiane Eilders. 2017. Design Matters! An Empirical Analysis of Online Deliberation on Different News Platforms: Design of Online Deliberation Platforms. *Policy Internet* 9, 3 (September 2017), 321–342. DOI:<https://doi.org/10.1002/poi3.154>
- [18] Rune Karlsen, Kari Steen-Johnsen, Dag Wollebæk, and Bernard Enjolras. 2017. Echo chamber and trench warfare dynamics in online debates. *Eur. J. Commun.* 32, 3 (June 2017), 257–273. DOI:<https://doi.org/10.1177/0267323117695734>

- [19] Travis Kriplean, Jonathan Morgan, Deen Freelon, Alan Borning, and Lance Bennett. 2012. Supporting reflective public thought with considerit. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, ACM, Seattle Washington USA, 265–274. DOI:<https://doi.org/10.1145/2145204.2145249>
- [20] Mariana L. De Carvalho, Thomas Olsson, and Joel Kiskola. 2021. Exploration of user interface mechanisms with affect labeling to enhance online discussion. In *Academic Mindtrek 2021*, ACM, Tampere/Virtual Finland, 210–220. DOI:<https://doi.org/10.1145/3464327.3464332>
- [21] Jinfen Li and Lu Xiao. 2020. Emotions in online debates: Tales from 4FORUMS and CONVINCEME. *Proc. Assoc. Inf. Sci. Technol.* 57, 1 (October 2020). DOI:<https://doi.org/10.1002/pra2.255>
- [22] Lisa B. Limeri, Nathan T. Carter, Jun Choe, Hannah G. Harper, Hannah R. Martin, Annaleigh Benton, and Erin L. Dolan. 2020. Growing a growth mindset: characterizing how and why undergraduate students’ mindsets change. *Int. J. STEM Educ.* 7, 1 (December 2020), 35. DOI:<https://doi.org/10.1186/s40594-020-00227-2>
- [23] Maciek Lipinski-Harten and Romin W. Tafarodi. 2013. Attitude moderation: A comparison of online chat and face-to-face conversation. *Comput. Hum. Behav.* 29, 6 (November 2013), 2490–2493. DOI:<https://doi.org/10.1016/j.chb.2013.06.004>
- [24] Farshid Marbouti and Alyssa Friend Wise. 2016. Starburst: a new graphical interface to support purposeful attention to others’ posts in online discussions. *Educ. Technol. Res. Dev.* 64, 1 (February 2016), 87–113. DOI:<https://doi.org/10.1007/s11423-015-9400-y>
- [25] Albert. Mehrabian. 1971. *Silent messages*. Wadsworth Pub. Co., Belmont, Calif.,.

- [26] Parviz Molavi, Niloofar Mikaeili, Mohammad Ali Ghaseminejad, Zhila Kazemi, and Misagh Pourdonya. 2018. Social Anxiety and Benign and Toxic Online Self-Disclosures: An Investigation Into the Role of Rejection Sensitivity, Self-Regulation, and Internet Addiction in College Students. *J. Nerv. Ment. Dis.* 206, 8 (August 2018), 598–605. DOI:<https://doi.org/10.1097/NMD.0000000000000855>
- [27] Raymond S Nickerson. Confirmation Bias: A Ubiquitous Phenomenon in Many Guises.
- [28] Alisa Rieger, Qurat-UI-Ain Shaheen, Carles Sierra, Mariet Theune, and Nava Tintarev. 2022. Towards Healthy Engagement with Online Debates: An Investigation of Debate Summaries and Personalized Persuasive Suggestions. In *Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*, ACM, Barcelona Spain, 192–199. DOI:<https://doi.org/10.1145/3511047.3537692>
- [29] Eric Schwitzgebel. 2019. Introspection. In *The Stanford Encyclopedia of Philosophy* (Winter 2019), Edward N. Zalta (ed.). Metaphysics Research Lab, Stanford University. Retrieved May 29, 2023 from <https://plato.stanford.edu/archives/win2019/entries/introspection/>
- [30] Bryan Semaan, Heather Faucett, Scott P. Robertson, Misa Maruyama, and Sara Douglas. 2015. Designing Political Deliberation Environments to Support Interactions in the Public Sphere. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ACM, Seoul Republic of Korea, 3167–3176. DOI:<https://doi.org/10.1145/2702123.2702403>
- [31] B. Shneiderman. 1996. The eyes have it: a task by data type taxonomy for information visualizations. In *Proceedings 1996 IEEE Symposium on Visual Languages*, IEEE Comput. Soc. Press, Boulder, CO, USA, 336–343. DOI:<https://doi.org/10.1109/VL.1996.545307>

- [32] Nadia Siddiqui and Enamul Hoque. 2020. ConVisQA: A Natural Language Interface for Visually Exploring Online Conversations. In *2020 24th International Conference Information Visualisation (IV)*, IEEE, Melbourne, Australia, 440–447.
DOI:<https://doi.org/10.1109/IV51561.2020.00077>
- [33] Paul Slovic, Melissa L. Finucane, Ellen Peters, and Donald G. MacGregor. 2007. The affect heuristic. *Eur. J. Oper. Res.* 177, 3 (March 2007), 1333–1352.
DOI:<https://doi.org/10.1016/j.ejor.2005.04.006>
- [34] Abhay Sukumaran, Stephanie Vezich, Melanie McHugh, and Clifford Nass. 2011. Normative influences on thoughtful online participation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, Vancouver BC Canada, 3401–3410. DOI:<https://doi.org/10.1145/1978942.1979450>
- [35] John Suler. The Online Disinhibition Effect.
- [36] H Tajfel. 1982. Social Psychology of Intergroup Relations. *Annu. Rev. Psychol.* 33, 1 (January 1982), 1–39. DOI:<https://doi.org/10.1146/annurev.ps.33.020182.000245>
- [37] Samuel Hardman Taylor, Dominic DiFranzo, Yoon Hyung Choi, Shruti Sannon, and Natalya N. Bazarova. 2019. Accountability and Empathy by Design: Encouraging Bystander Intervention to Cyberbullying on Social Media. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW (November 2019), 1–26. DOI:<https://doi.org/10.1145/3359220>
- [38] Adhika Pramita Widyassari, Supriadi Rustad, Guruh Fajar Shidik, Edi Noersasongko, Abdul Syukur, Affandy Affandy, and De Rosal Ignatius Moses Setiadi. 2022. Review of automatic text summarization techniques & methods. *J. King Saud Univ. - Comput. Inf. Sci.* 34, 4 (April 2022), 1029–1046. DOI:<https://doi.org/10.1016/j.jksuci.2020.05.006>

- [39] Scott Wright. The Role of the Moderator: Problems and Possibilities for Government-Run Online Discussion Forums.
- [40] Amy X. Zhang and Justin Cranshaw. 2018. Making Sense of Group Chat through Collaborative Tagging and Summarization. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW (November 2018), 1–27. DOI:<https://doi.org/10.1145/3274465>
- [41] Amy X. Zhang, Lea Verou, and David Karger. 2017. Wikum: Bridging Discussion Forums and Wikis Using Recursive Summarization. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, ACM, Portland Oregon USA, 2082–2096. DOI:<https://doi.org/10.1145/2998181.2998235>
- [42] 2023. What is karma? *Reddit Help*. Retrieved May 8, 2023 from <https://support.reddithelp.com/hc/en-us/articles/204511829-What-is-karma->
- [43] r/changemyview. *reddit*. Retrieved May 10, 2023 from <https://www.reddit.com/r/changemyview/>
- [44] Debate platform powered by reason - A tour of Kialo. *Kialo*. Retrieved May 18, 2023 from <https://www.kialo.com/tour>
- [45] YouTube. Retrieved June 27, 2023 from <https://www.youtube.com/>
- [46] About Community Notes on Twitter | Twitter Help. Retrieved June 1, 2023 from <https://help.twitter.com/en/using-twitter/community-notes>
- [47] GPT-4. Retrieved June 1, 2023 from <https://openai.com/product/gpt-4>

