

# Making Online Communities ‘Better’: A Taxonomy of Community Values on Reddit

Galen Weld, Amy X. Zhang, Tim Althoff

Paul G. Allen School of Computer Science & Engineering, University of Washington  
{gweld, axz, althoff}@cs.washington.edu

## Abstract

Many researchers studying online communities seek to make them better. However, beyond a small set of widely-held values, such as combating misinformation and abuse, determining what ‘better’ means can be challenging, as community members may disagree, values may be in conflict, and different communities may have differing preferences as a whole. In this work, we present the first study that elicits values directly from members across a diverse set of communities. We survey 212 members of 627 unique subreddits and ask them to describe their values for their communities in their own words. Through iterative categorization of 1,481 responses, we develop and validate a comprehensive taxonomy of community values, consisting of 29 subcategories within nine top-level categories, enabling principled, quantitative study of community values by researchers. Using our taxonomy, we reframe existing research problems, such as managing influxes of new members, as tensions between different values, and we identify understudied values, such as those regarding content quality and community size. We call for greater attention to vulnerable community members’ values, and we make our codebook public for use in future research.

## 1 Introduction

Online communities account for an ever-increasing share of all human interaction (Williamson 2020). People use millions of different online social communities for accessing news (Weld, Glenski, and Althoff 2021), for entertainment, and for socialization, amongst many other purposes. Unfortunately, these online communities have been shown to have substantial harms to both community members and society as a whole, including the distribution of misinformation, harassment and bullying, and coordinated activity to undermine elections. As a result, many researchers are working on methods to understand these harms and make online communities ‘better.’ However, truly understanding how to make communities ‘better’ requires going beyond simply mitigating and minimizing harms and going towards an understanding of community *values*. Determining a community’s values is a non-trivial challenge, as communities have many stakeholders with divergent preferences.

The term ‘values’ can have many meanings; here we use the definition from Value Sensitive Design: ‘what a person

or group of people consider important [to their online community]’ (Friedman, Kahn, and Borning 2006, pg.2). In this definition, values not only have a *topic*, *e.g.*, the diversity of the community, but also a *preference*, *e.g.*, a preference for *more* diversity. Value Sensitive Design is a design framework which underpins our work. At its core, Value Sensitive Design suggests that values must be considered when designing in contexts such as online communities (Friedman, Kahn, and Borning 2006).

Understanding community values is challenging because values can vary widely both between and within communities, and values can conflict with one another. A community focused on mental health support may want to foster inclusion more than a community focused on financial trading, a difference in values between communities. Within the same community, members may have different value preferences regarding the same value topic, *e.g.*, one community member may desire more diversity, while another may wish the community was more homogeneous. Someone who is a member of two (or more) communities may even have different preferences in different communities; that same person may prefer more diversity for one community and more homogeneity for the other. Lastly, values of different topics may conflict with one another as well. For example, while an online photography discussion community may desire to create a space that is welcoming to beginners, this value conflicts with the same community’s desire to hear particularly from expert photographers who may be perceived as having the most to contribute to the conversation. These differences and conflicts are a critical consideration for researchers as well as community moderators and members.

Although much has been published on positive aspects of online communities (Robert E Kraut et al. 2012), previous work which seeks to make online communities better has largely focused on specific aspects of online communities, especially those which are widely agreed to be harmful, such as harassment, rule-breaking, and misinformation. However, upon deeper inquiry, even these more commonly studied harms are quite complex, with substantial disagreement within and across communities regarding the extent to which these ostensibly harmful behaviors should be tolerated (Jiang et al. 2021; Scheuerman et al. 2021).

Implicit in much of this prior work is an assumption of communities’ and their members’ values, yet exactly what

values communities hold has not yet been comprehensively studied. While we do not argue that any one set of values is superior to others, we believe that a critical first step towards improving online communities is developing a comprehensive understanding of community members' own values, not just obvious ones that researchers have assumed all communities care about.

In this work, we survey redditors to answer the research question “What values do community members hold for their communities?” Through a series of advertisements placed on reddit, we recruit 212 people and collect 1,481 free response answers to questions about the values they hold for the 627 unique communities they consider themselves a part of (§2). To the best of our knowledge, this is the first such survey to gather community members' values in their own words.

We apply an iterative categorization methodology to produce our primary contribution: a taxonomy of community values with nine top-level categories and 29 subcategories (Table 1). We validate this taxonomy with a held-out set of 1,180 additional responses to demonstrate saturation, and achieve very high inter-rater agreement (Fleiss kappa = 0.874) using a codebook which we make public to support future research (Appendix E). Among the other key findings we contribute, we find that online social community members' values cover a broad range of categories from technical features to the diversity of community members, and that the quality of content submitted to communities is the most frequently reported value category (§3).

Our work enables new research to improve online communities (§4). In addition to working on what they consider valuable, researchers seeking to improve online communities should work on values we found to be frequently held yet understudied. Open research challenges with regard to understudied values include measuring the quality of content in the specific context of different communities, and the difficulty of growing communities' membership while maintaining meaningful community interaction (§4.1). Now that we have delineated a taxonomy of values, we can also begin to examine how values relate to each other. We argue that some widely held community values are inherently in tension, such as maintaining the quality of content in a community while simultaneously including new members, and call for research into community design with these tensions in mind (§4.2). Finally, we call for additional work on reconciling differences in values within communities, including community governance that protects the values of vulnerable community members (§4.3).

## 2 Survey and Categorization Methodology

### 2.1 Reddit Background & Context

In this work, we focus on reddit. reddit is an ideal platform for conducting research on the values of online communities, as (unlike other social media platforms such as Twitter) reddit is explicitly divided into thousands of unique communities, known as ‘subreddits,’ each with their own topical foci, rules, moderators, norms, and enforcement practices. Furthermore, almost all content on reddit is publicly avail-

able (Baumgartner et al. 2020), and reddit has been widely examined by the research community (Medvedev, Lambliotte, and Delvenne 2019). On reddit, subreddit’s names are prefixed with /r/, and we adopt this notation in the rest of this paper.

### 2.2 Survey Instrument

All responses were gathered through an online survey hosted on the Qualtrics platform. The survey is summarized here and included in its entirety in Appendix D. The survey consists of five sections: (1) informed consent, (2) demographic questions, (3) general questions about the respondent’s usage of reddit, (4) subreddit specific questions, and (5) reflection questions. All questions except the informed consent question are optional.

Before any other questions are posed, the participant is shown a brief summary of the nature and aims of the survey, along with IRB information (for more details, see §2.7) and is asked for their consent. Then, in the demographics section, the participant is asked to describe their age, gender identity, and racial identity. Due to the challenge of obtaining parental consent over the internet, minors (people under the age of 18 in our jurisdiction) are excluded from participating.

Next, the participant is asked to optionally provide their reddit username, which is used to query the reddit API for a list of subreddits in which the participant recently posted or commented, and as a contact point for the raffle (§2.3).

Once the general reddit questions are answered, the participant is shown a list of the five most recent subreddits they posted or commented in (Appendix A), and asked to remove any subreddits that they do not consider themselves a member of, and to add any subreddits they consider themselves a member of that are missing from the list. Participants who decline to provide their username or whose username was not found on reddit are presented with an empty list for them to populate themselves, and entered subreddits are checked in real time against the reddit API to preclude spelling errors and to ensure the subreddit exists.

The next section is dedicated to subreddit specific questions. It consists of two free-response questions which are asked in turn for every subreddit listed by the participant in the previous step: *As it exists right now*, what are a few of the best aspects of the /r/<subreddit>community? *If you could change anything*, what are some aspects of the /r/<subreddit>community you would like to improve upon? We chose to use these two questions to elicit participants’ values from their specific, real experiences, rather than asking participants to speculate about abstract cases. To this goal, we chose to ask about positive aspects of the community as well as negative aspects, as these two questions elicit different values: the positive question offers insight into participants’ values which are well-implemented by their communities, whereas responses to the negative question show which of the participants’ values are less accommodated by current community practices. Asking both of these questions allows us to gather the broadest possible set of values for categorization, crucial to developing a comprehensive understanding of community values.

The survey was piloted with 13 participants from several departments in two large American universities. All 13 pilot participants denied having any difficulty understanding any of the survey questions or interacting with the online survey tool.

### 2.3 Participant Recruiting and Incentives

The 212 participants in this study were recruited primarily through reddit advertisements. These advertisements display inline with other content in both individual subreddits and aggregated content views, and are shown on the reddit website as well as the official reddit mobile app. Appendix A Figure 2 shows the appearance of these advertisements. To increase the diversity of recruiting, the survey was also distributed to relevant university mailing lists and Slack channels at two large American universities, as well as posted to /r/SampleSize, a subreddit for the distribution of surveys, recruiting an additional 41 participants. Participation was incentivized with a raffle. Additional details on recruiting pipeline attrition and the raffle are located in Appendix B.

### 2.4 Participant Demographics

In general, we find that the demographics of our 212 survey respondents match overall platform demographics closely (Barthiel et al. 2016) (summary of respondent demographics in Appendix C). As with reddit's overall userbase, our respondents skew young, white, and more commonly identify as men. Compared to reddit's overall demographics, the age our of respondents is similar, with 71% of our respondents being under the age of thirty. People who identify as non-binary or do not identify with man or woman are slightly overrepresented in our survey results compared to reddit as a whole. When it comes to racial and ethnic identities, white and Black users are slightly underrepresented in our responses, while Hispanic and Asian responses are slightly overrepresented.

### 2.5 Iterative Categorization

Once the survey responses were collected, all free-text responses were divided into idea units (Strauss 1987), where each idea unit represents a distinct thought. For example, the response 'The content is educational and I like how the community is engaged' would be divided into two idea units: 'The content is educational' and 'I like how the community is engaged.'

Our taxonomy was produced using an initial set of 301 idea units gathered from the first 39 respondents to the survey, with the remaining 1,180 idea units held out for validation. Using a grounded theory approach (Glaser, Strauss, and Strutz 1968), a team of five researchers worked together to iteratively categorize the initial idea units using an inductive coding method (MacQueen et al. 1998). While respondents often expressed value preferences in their responses, only value topics were used to categorize idea units. The researchers worked independently to initially cluster similar idea units, then came together to resolve differences in clustering until a consensus was reached. The initial tentative clusters were assigned names and definitions to produce

a working taxonomy, then the researchers collaboratively re-categorized all idea units, creating and removing categories under group consensus. This process was repeated three additional times until the iterative process converged and no further changes were needed to satisfactorily categorize all idea units. Once this was completed, the researchers worked together to write a codebook (Appendix E) describing the taxonomy, which is hierarchical, with top-level categories and subcategories. When possible, idea units are assigned to the more specific subcategories, with top-level categories reserved for broad or vague idea units.

### 2.6 Inter-Rater Reliability and Validation of the Taxonomy

To validate the codebook, three researchers were trained on the codebook and independently labeled 100 idea units randomly sampled from the held-out set. Inter-rater reliability was very high, with a Fleiss kappa of 0.874 when considering all 29 subcategories. When measuring agreement on only the nine top-level categories, the three raters were in even greater agreement (Fleiss kappa = 0.902). This level of agreement would typically be described as "almost perfect" (Landis and Koch 1977) or "excellent" (Gwet 2014), and demonstrates that the taxonomy categorization using our codebook is robust and largely unambiguous.

To validate the taxonomy, the 1,180 held-out idea units were coded by a single researcher. Every single one of the idea units in the validation set was able to be categorized using the taxonomy codebook, demonstrating saturation.

### 2.7 Ethical Considerations & Broader Impacts

In order to ensure the anonymity of our participants, we keep all their responses confidential. Access to responses was limited to only the immediate research team, and all responses quoted in the following sections have been paraphrased to ensure participant anonymity. All data collected in this study was provided by participants who were informed of the nature of the study, the potential risks of participation, and who consented to participate. We do not make any use of participants' public reddit histories for any purpose, as the use of such data has been shown to make many participants uncomfortable (Fiesler and Proferes 2018). This research was reviewed and approved by the University of Washington IRB under ID STUDY00011457

## 3 Taxonomy of Community Values

Our taxonomy has two hierarchical levels—the lower level consisting of 29 specific subcategories grouped into nine top-level categories. These categories reflect the diverse range of value topics reported by our participants, ranging from *Technical Features* to *Diversity* (Table 1). Categories are grouped according to value topic, regardless of value preference, and thus may contain conflicting value preferences, e.g., the *Size* category contains idea units from respondents who prefer both larger and smaller sizes for their respective communities. *Quality of Content* is the most frequently mentioned value category, with 47% of all idea

	Subcategory Name	Example Idea Unit	Frequency
Quality of Content	Quality of Content	"I like the content in /r/HistoryMemes"	31
	a. Personal Preferences	"[r/StallmanWasRight] has too many articles and not enough personal stories"	293
	b. Education, Entertainment	"it's interesting to learn about other countries [on /r/AskEurope]"	247
	c. Curation, Recency, Discovery	"[I like that] I can keep up to date with new releases on /r/hiphopheads"	60
Community Engagement	d. Spam, Reposts, Bots	"[There are] too many reposts of things people [on /r/Embroidery]"	46
	Community Engagement	"the connection between users is great [on /r/vexillology]"	50
	a. Quality of Interaction or Community as a Whole	"[I like that] people [on /r/gaidhilig] are quick to respond with helpful comments"	170
Size	b. Connection, Universalization	"being on /r/teenagers shows me that I am not alone in what I am going through"	55
	Size	"[I like how] /r/stocks is big"	2
	a. Volume of Content	"because [/r/AskReddit] is so large, lots of interesting questions come up"	76
Participation & Inclusion	b. Size of Community	"[I like how /r/roguelikedev is a] tiny community with a clear purpose and scope"	36
	Participation & Inclusion	"[I like that /r/CrossStitch] is so inclusive"	16
	a. Offensive, Abusive, Harassing Content or Behaviors	"I don't like some of the sexist jokes [on /r/ProgrammerHumor]"	58
Diversity	b. Outsiders, Demographics, Limits	"[I think] too many people from outside Seattle comment on posts [in /r/Seattle]"	12
	c. Tools for Participation	"[/r/snails] should add an easily accessible 'beginners' questions' section"	7
	Diversity	"[I like that /r/askscience is so] diverse"	1
a. Variety of Content	a. Variety of Content	"[I like how] I get to see different isopods [in /r/isopods], both wild and captive"	56
b. Diversity of People	b. Diversity of People	"[/r/AskCulinary is] a nice mix of pros, experienced home cooks, and newcomers"	19
Mods	Moderation & Moderators	"[I like how] mods of /r/goodanimememes don't change rules without a vote"	69
Norms	Norms	"[I dislike that] /r/alberta has weird rules"	3
	a. Adherence to Norms	"[I don't like that] people don't read the [/r/whatisthisplant] FAQ before posting"	46
	b. Voting Behavior	"I wish people [on /r/ACMilan] wouldn't use downvote as a disagree button"	10
Technical Features	Technical Features	"[It frustrates me that] posts [on /r/skiing] never load properly"	27
	a. Flair, Tags, NSFW labels	"flairs [on /r/ImmigrationCanada] are very useful here"	12
	b. Search, Filters	"I wish I could filter by genre of game [on /r/ShouldIbuythisgame]"	11
Trust	c. Recommendation Systems	"I would like more recommendation measures [on /r/lotrmemes]"	3
	Trust		0
	a. Knowledgeable People	"[I like how /r/eyetriage has] real doctors in the community, so it's more reliable"	22
b. Trustworthy Content	b. Trustworthy Content	"[I like how /r/nyu] is legit because students are honest in their opinions"	18

Table 1: Summary of our taxonomy of values, along with example quotes from participants. Although the table is sorted by frequency, we note that the most frequent values are not necessarily the most important.

units, while *Technical Features* and *Trust* are the least frequently mentioned categories, each with less than 4% of the idea units.

While we report counts for the number of idea units falling into each category and subcategory, we do not make quantitative claims or comparisons between different categories or different communities. Communities, each with their own subject matter, rules, and membership, are different from one another, which is why our survey is designed to ask participants about each community individually. Rather than quantify the importance of any one category, our goal and contribution is to understand the broad diversity of value topics across communities in greater detail. The following subsections describe each taxonomy category in detail.

### 3.1 Quality of Content

The *Quality of Content* category is the largest of the categories that compose our taxonomy, containing 4 subcategories and 46.5% (677/1481) of all idea units derived from responses. It is unsurprising that most community members would have many values relating to the content of the communities they are a part of, as in most reddit communities, sharing content is the primary mode of interaction with the community. We expect values in this category to be especially prevalent in communities that are less focused on engagement and connection with others and most focused on content itself (e.g., meme sharing communities).

The four subcategories in this category (*Personal Preferences*, *Education/Entertainment*, *Curation/Recency/Discovery*, and *Spam/Reposts/Bots*) reflect the broad range of aspects of content and its presentation

that are liked or disliked by community members. Many community members appreciate the curation and discovery of content that is provided by the community through mechanisms such as up- and down-voting, *e.g.*, “I keep up to date with new releases on /r/hiphopheads.” The most frequently mentioned subcategory within the top-level *Quality of Content* category is *Personal Preferences*, which reflects how well (or poorly) the content aligns with personal preferences for specific types of content, such as memes (“the best part of /r/Terraria is the memes”) or different subjects (“[in /r/BalticStates,] the posts about Estonia are the best because it’s such a fabulous country”). These *Personal Preferences* are difficult to generalize because in most cases they are specific to the subject of the community as a whole.

Idea units regarding *Education/Entertainment* and *Spam/Reposts/Bots* tend to be more consistent from subreddit to subreddit. In general, content that is educational or entertaining is especially liked, *e.g.*, “/r/knitting teaches me new stitches and patterns”, while *Spam/Reposts/Bots* are fairly universally disliked, and manifest more similarly in different communities. However, the exact nature of reposts and repeated content can vary somewhat from subreddit to subreddit, *e.g.*, “I wish users [of /r/AskReddit] would quit asking the same questions over and over just to get karma”<sup>1</sup>.

### 3.2 Community Engagement & Interaction

The *Community Engagement* category contains 18.9% (275/1481) of the idea units derived from responses, making it the second most frequently reported value topic amongst respondents. We divide these idea units into two subcategories: *Quality of Interaction or of the Community as a Whole*, and *Connection and Universalization*, *i.e.*, the realization that others exist with similar interests/feelings. The vast majority (256/275) of these responses were mentioned in a positive context, suggesting that redditors mostly feel positively about community engagement.

Comments on *Quality of Interaction or of the Community as a Whole* mentioned both qualities of the individual interactions with community members (*e.g.*, “I often ask for help with language learning resources [on /r/gaidhlig], and everyone is always quick to respond”) as well as qualities of the community as a whole (*e.g.*, “I appreciate the goodhearted nature of most of the people [on /r/Konosuba, a community dedicated to the eponymous Japanese novel series].”) However, 19 participants were unhappy with the quality of the interactions in their communities, such as “[On /r/msu] there are often gatekeepers who comment on posts. However, these people tend to get downvoted quickly, so it’s not that big of a deal.” This community member then suggested that “a brief note about negativity in the sub’s rules could help with this.”

*Connection and Universalization* is especially often-mentioned in communities for people with a common identity or interest who may be physically far apart or part of a minority group and therefore unable to connect as easily offline. One respondent wrote “[I love that /r/blackladies] is

a community of black women coming together to discuss social issues that are prevalent and important to us.” Another says “[/r/Glaucoma] lets me get in touch with people around the world [who are] dealing with a similar health issue.” These quotes from participants reflect a body of literature that finds that online communities can be helpful venues for minority groups and those with special needs connect with similar people for support (Raj and Daniels 2017; Kamalpour, Watson, and Buys 2020).

### 3.3 Size

Many responses (144/1481) commented on the *Size* of communities, both regarding the *Volume of Content* submitted to the community and the number of people in that community (*Size of Community*). Participants were varied in their preferences, with some preferring a larger volume of content (*e.g.*, “I like that /r/assholedesign is regularly updated”) and others preferring less content. One member of /r/SampleSize, a community for distributing surveys, wrote “The large volume of posts means that you need to time your submission very carefully in order to make sure that your survey doesn’t get buried.”

Similarly, respondents were fairly evenly split (21 preferring larger, 15 preferring smaller) on their desire for larger or smaller communities. Those preferring larger communities perceived those larger communities as offering increased opportunities for interaction (“More active users [on /r/photocritique] would make it easier to engage with like minded individuals through their posts.”), while those preferring smaller communities tended to like the specificity of subject matter and focus, *e.g.*, “[I like that /r/roguelikedev, a community dedicated to the development of a specific type of video game, is a] tiny community with a very clear purpose and scope.” Community members’ value preferences regarding community size seem likely to depend on many factors, especially including the subject of the community and the nature of interactions which occur within.

### 3.4 Participation & Inclusion

*Participation and Inclusion* in the community was a value topic reported by 93 respondents. We divide their responses into three subcategories: *Tools for Participation*, *Offensive/Abusive/Harassing Content or Behaviors*, and *Outsiders/Demographics/Limits*, which focuses on *who* participates.

Comments on *Offensive/Abusive/Harassing Content or Behaviors* were split (20 units to 38 units) between positive comments (praising the absence of offensive content), *e.g.*, “[I like that /r/WANDAVISION] has no homophobia, racism, or any discrimination,” and comments that described the respondent’s experience with such behavior that was detrimental to inclusion. One member of /r/sewing wrote “sometimes the commenters on a post will write personal things about a poster’s appearance, which I don’t think is appropriate.”

Community members who commented on *who* participates in their communities (*Outsiders/Demographics/Limits*) were frequently concerned with the presence of outsiders, who are perceived as not belonging

<sup>1</sup>Karma is reddit’s name for points gathered through the receipt of upvotes.

to the community by virtue of their lack of familiarity with the subject or even their physical location. One member of /r/Seattle wrote that “too many people from outside Seattle comment on posts in this subreddit.” In any community with a specific subject, some degree of boundaries for membership are natural, yet too much insularity in online communities can be harmful (Allison and Bussey 2020).

*Tools for Participation* were suggested by seven participants who felt their communities were lacking such tools. One participant wrote “[/r/snails] should add an easily accessible ‘beginners’ questions’ section.” Some tools, such as automated posts and messages (Yazdanian et al. 2019) and badges, have been found to reduce unwelcoming reactions to new community members (Santos et al. 2020) as well as such members’ compliance with rules (Matias 2019b).

### 3.5 Diversity

Many community members commented on the *Diversity* of their communities, which we divide into two subcategories: *Variety of Content* and *Diversity of People*. More respondents (56/75) commented on *Variety of Content*, e.g., “[/r/CollegeBasketball] has an ideal blend of banter, rumors, statistical insights, and glorious shitposting.” Those who commented on *Diversity of People* more frequently commented on aspects of the community they would like to change, e.g., “I wish [/r/knitting] had more variety in skill. Right now it’s mostly skilled knitters, whereas it would be appreciated to see some beginner knitters.”

### 3.6 Moderation & Moderators

*Moderation & Moderators* are controversial subjects on reddit, and many moderators feel strongly disliked by members of the community they moderate (Matias 2019a). However, close to half (31/69) of idea units regarding moderation were positive, with respondents praising the moderator team in general (“mods [of /r/Phillipines] have done an excellent job of maintaining the community”) as well as specific moderators (“<username redacted> is such a great mod [of /r/ApplyingToCollege] who is super helpful.”).

Of the 38 negative comments on moderation, many of them were critical of perceived arbitrary rule enforcement, e.g., “moderators arbitrarily remove posts because they’re ‘against the subreddit’s rules.’” Other idea units requested more active moderation (14 units), complained about perceived power abuses (8 units), and called for greater community involvement in rule making (2 units).

### 3.7 Norms, Voting Behavior, and Adherence

Almost every (54/59) response regarding *Norms* was on a negative aspect of the community, suggesting that norms are mostly noticeable in a community when they are violated. Many of the idea units relating to norms (46/59) were community specific and focused on *Adherence to Norms*, such as newcomers not reading the FAQ before posting, posters not providing adequate information and expecting a response (/r/whatisthisplant), or not including appropriate sources. Many respondents also requested additional rules or changes to norms (19 units), e.g., “I would like to reduce the

amount of people speaking English [on /r/ich\_iel, a German-speaking subreddit].”

The remainder of idea units (10/59) were complaints about *Voting Behavior*, particularly the use of the “downvote button as a disagree button.”

### 3.8 Technical Features

53 idea units (3.6%) focused on *Technical Features*, with three subcategories: *Flairs/Tags/NSFW<sup>2</sup> Labels, Search and Filters*, and *Recommendation Systems*. Most (36/53) idea units in *Technical Features* focused on negative aspects of technical features, particularly such features’ absence or failure to work properly. Most positive idea units praised their respective subreddits’ use of flair and tags, e.g., “NSFW tagging is perfectly used here [in /r/hemorrhoid].”

While flairs, tags, recommendations systems, filtering options, and quality search functionality can dramatically improve users’ experiences in online communities, on reddit, such technical features can be difficult for communities to implement and modify, as they are controlled by the reddit administration, not community members. Frequently, incentives are misaligned between subreddit leadership and the reddit administration, who may be unwilling to implement new technical features at the request of communities. As a workaround, many communities implement their own second-party technical features by modifying or repurposing existing features, such as custom community moderation bots (Kiene and Hill 2020) or reputation systems such those in /r/changemyview (Jhaver, Vora, and Bruckman 2017).

### 3.9 Trust

The *Trust* of a community was the least frequently commented on value category. This category consists of two subcategories, *Knowledgeable People* and *Trustworthy Content*, again differentiated by a focus on people vs. content. Of respondents who commented on the trustworthiness of people, one respondent appreciated the credentials of community members “There are many doctors [on /r/eyetriage] so it’s a lot more reliable,” whereas another respondent lamented how anonymity interfered with the trustworthiness of the /r/MachineLearning community: “Anonymity sometimes makes it so I don’t know who is qualified to say what.”

## 4 Implications & Discussion

Community members value a broad range of topics for their communities ranging from *Diversity* (§3.5) to *Technical Features* (§3.8). However, some of these topics, such as the *Quality of Content* (§3.1), are much more frequently reported by community members than others. We do not, however, suggest that these topics which are most frequently mentioned should be considered more important to *researchers* than less frequently mentioned topics, such as *Trust* (§3.9). Most existing research that seeks to make online communities ‘better’ focuses on implicit community values which are relatively narrow in scope, especially those with broader societal impact (e.g., mis/disinformation and

<sup>2</sup>A common reddit acronym meaning ‘not safe for work’ and used to indicate content containing gore or nudity.

political polarization) and/or a negative impact on vulnerable populations (e.g., abuse, harassment) (§5). While these research directions are critical to mitigating major harms associated with online communities, our findings suggest that many values held by community members are understudied, complex, and often in conflict with one another. In this section we discuss this complexity (§4.1) and conflict (§4.2), implications for moderation and governance practices (§4.3), and subsequently how our taxonomy relates to other taxonomies from different contexts (§5).

## 4.1 Understudied Community Values

Many values explicitly stated by community members have topics which are not well studied by the research community. We find that *Quality of Content* is the most frequently mentioned value topic (§3.3), accounting for 46.5% of idea units in our responses, yet with the exception of spam and bot detection, this value is poorly understood. *Quality of content* is arguably the most challenging value topic to define and measure, as it is extraordinarily context-specific and dependent on the nature of the community in question. For example, content that is high quality in */r/catpictures* would be woefully inappropriate for */r/science*. As existing work in this space focuses mostly on understanding the quality of conversational (and other text-based) content (Zhang et al. 2018b,a; Lakkaraju, McAuley, and Leskovec 2013), additional work which contributes to the understanding of the context-specific quality of image and video-based content is especially needed.

Community *Size* is another value category that is understudied and complex. Some work has studied the impact of rapid growth on communities (Lin et al. 2017; Kiene, Monroy-Hernández, and Hill 2016); however, our findings suggest that members perceive a difference between changes in the size of the community (e.g., the number of participants) and the volume of content (§3.3). While most respondents (47/76) prefer a larger volume of content, many respondents (15/36) perceive smaller communities as better due to having stronger community engagement and interactions, e.g., “as the community has grown [it] has lost its small and friendly community feel”. Balancing this tension is an important area for future research, as is understanding how desired community size varies as a function of the community member’s relationship with the community in question.

## 4.2 Conflicting Community Values

Our results show that community members hold a broad range of values, but these values are challenging to implement because values in different categories often conflict with one another.

**Inclusion vs. Quality of Content and Norms.** While the challenge of incorporating new members has been previously identified and studied (Robert E Kraut et al. 2012; Danescu-Niculescu-Mizil et al. 2013; Cho and Wash 2021), our findings permit the framing of this tension as a conflict between *Inclusion* (of new members) and *Quality of Content and Norms*. One participant expressed this sentiment in their

response, saying “It’s frustrating when [new members] tend to leave out information they’re expected to include.” In social media communities, some work has experimented with onboarding documents and mentorship (Santos et al. 2020; Matias 2019b; Yazdanian et al. 2019) to improve new members’ understanding of community norms and maintain the quality of content. Methods to mitigate the tension between inclusion, norms, and quality of content have been studied more deeply in the context of peer production communities such as Wikipedia (Ciampaglia and Taraborelli 2015; Halfaker et al. 2013; Halfaker, Kittur, and Riedl 2011), and there is great potential for future work to study how these findings generalize to more social communities and develop new tools (Robert E Kraut et al. 2012, Ch.5). To an extent, however, these values are inherently at odds.

**Size vs. Community Engagement.** Community engagement and size are also often in conflict with one another. While several studies have found that, by some metrics, communities’ health is not harmed in the long term by increases in size (Lin et al. 2017; Kiene, Monroy-Hernández, and Hill 2016), growth is a frequent subject of complaint across many platforms. One of our participants wrote “[/r/formula1 is] such a large community that is hard to engage with other members.” On the other hand, many participants also reported desiring more activity and more frequent content in their communities. An open and important research question is how to scale communities to larger size without sacrificing the sense of interpersonal engagement.

## 4.3 Community Governance & (Lack of) Consensus

While the previous section described conflicts between different value categories, we also find evidence that members of the same community can disagree with one another’s value *preferences* for the same value category. For example, eight participants who are members of */r/AskReddit* expressed differing preferences for what content is permitted in the community. Three idea units indicated a preference for an ‘anything goes’ strategy, while the remaining four wished for more restrictions on specific subjects, mainly sex and drug use. Other research has found similar evidence that community members often disagree with one another on matters such as the severity of harmful behavior (Scheuerman et al. 2021; Jiang et al. 2021). Strategies for reconciling these differences of opinion, however, are not well understood. In the context of peer production communities such as Wikipedia, some research has already explored some sources of internal disagreement, such as tensions between senior and junior members (Halfaker, Kittur, and Riedl 2011; Halfaker et al. 2013; Steinmacher et al. 2015), however the extent to which these findings generalize to social media platforms such as reddit may be limited. Our results suggest that even relatively small samples of members are adequate to surface some differences in values between community members. Future work should examine the degree of agreement or disagreement on values held by many members from the same communities, and what factors are predictive of such differences in opinion.

**Affordances for Participatory Governance.** Participatory governance practices can help build consensus and reconcile differences in opinion when considering and implementing rule changes. However, affordances for such practices are extremely limited on social media platforms, where the vast majority of communities' rules are determined exclusively by a small set of moderators with no formal input from the broader community (Zhang, Hugh, and Bernstein 2020). We find (§3.6) that many community members perceive lack of transparency and arbitrary enforcement decisions as evidence of corruption ("mods will ban you without warning if you say something they disagree with"), and desire greater input into moderation ("moderators should consult the community about what we want"). Conflict between moderators and the broader community has been identified in some prior work (Matias 2019a), yet the exact differences between moderators' and nonmoderators' opinions have not yet been studied and quantified systematically. Increasing participatory governance practices in online communities may help alleviate some of this conflict, yet such practices are not a panacea, and can cause harm if not implemented carefully, *e.g.*, by increasing the burden of labor on certain groups, or by giving a veneer of legitimacy to unilateral decision making (Kelty 2017).

**Methods for Managing Irreconcilable Differences.** What happens when differences in values are too great to reconcile? On reddit, it is not uncommon for some members to splinter and create an alternative subreddit in response to perceived grievances or frustration with the status quo, however, to the best of our knowledge, this 'exit' phenomenon has not been studied in depth (Frey and Schneider 2021). Other communities, such as consensus-based peer production communities, have different practices to manage internal disputes such as formal arbitration committees (Konieczny 2017), but it is unclear how such practices would generalize to social media communities. In some cases of divergent values, such as differing *Personal Preferences* for content, personalized filtering may be used so that each user does not see content they wish to avoid (Jhaver et al. 2018). However, this undermines social translucence, a theory which suggests that making online behavior visible creates social spaces with shared accountability (Erickson and Kellogg 2000). Some research has explored interventions to balance the trade-off between social translucence and the personalization afforded by filtering (Gilbert 2012).

**Power Structures and Protecting Vulnerable Community Members.** Furthermore, it is likely that some of the categories of values we have identified in this work will be of special importance to vulnerable groups. For example, one member of /r/AskWomenOver30 said "everyone assumes I'm a white American, which really changes the dynamic when I ask about career, relationships, and more." In this case, only a small fraction of the community may recognize that these assumptions are harmful to the experience of some community members, but that minority perspective is still very important. If traditional participatory decision making practices such as voting were implemented naively, the majority (which may not have had personal experience with

harms such as the aforementioned assumptions of members' background, harassment (Matias, Simko, and Reddan 2020), *etc.*) may not support or even be opposed to measures designed to reduce these harms, which disproportionately affect women and other minority groups (Lenhart et al. 2016).

This phenomenon, known as 'Tyranny of the Majority,' is well known in the offline governance context (Guinier 1994), yet has not been studied in-depth in the online context. Some peer production communities, such as Wikipedia, emphasize consensus-based deliberation over vote-counting partly as a way to avoid this issue (Wikipedia 2021; Im et al. 2018), though issues with bias against women and other minority groups still persist (Tripodi 2021). Additional research into participatory governance practices for online communities which protect vulnerable groups is sorely needed.

At a higher level, much research on online communities studies these communities through the lens of moderators (Matias 2019a) and other people in positions of power (Robert E Kraut et al. 2012; Chandrasekharan et al. 2017; Shen and Rose 2019; Habib et al. 2019). This is especially the case for empirical work which relies upon rules and enforcement actions as concrete evidence of behavior and community norms (Jhaver et al. 2019; Jhaver, Bruckman, and Gilbert 2019; Chandrasekharan et al. 2018; Fiesler et al. 2018). While these rules and enforcement actions are natural sources of empirical data, researchers must take care and recognize that many community members feel as though they do not have an adequate stake in rulemaking and enforcement (§3.6). Online platforms' power structures are complex (Jhaver, Frey, and Zhang 2021), and while surveys such as this one are useful tools, additional work is needed on scalable methods for empirically studying community behavior and values that include the voices of the least empowered.

#### 4.4 Limitations

While we made intentional efforts to recruit participants from a diverse set of backgrounds by using multiple recruiting methods, it is possible that selection effects impacted who responded to our survey, which may have resulted in a taxonomy that is not truly representative of all community members. Furthermore, we do not include participants younger than the age of 18. Lastly, while our taxonomy derived from responses relating to a large set of 627 diverse communities, we only have responses from a small fraction of each community's membership. These facts, along with our (in absolute terms) low click-through rate of 0.227%, are indicative of a fundamental reality for researchers of online communities: it is very difficult to effectively poll an online community. Additional research is needed to improve polling methods, *e.g.*, by reducing friction in online surveys. Improved methods could strengthen results' representativeness and susceptibility to bias from sources such as poorly defined community membership.

We believe the validity of our taxonomy is demonstrated both by the large (1,180 idea unit) validation set used to demonstrate saturation (§2.5), the high inter-rater reliability (Fleiss kappa = 0.874) of the codebook, as well as the corre-

spondence between the values reported by our participants and those identified by sociologists and psychiatrists in the context of interpersonal interactions (§5). Making ‘truly representative’ samples of online communities’ membership is especially difficult given that such membership is often not well-defined, yet this is an important area for future work.

Our participants were asked about their experiences on a single social platform, reddit. While reddit is large, popular, and contains many thousands of communities covering diverse subjects (Medvedev, Lambiotte, and Delvenne 2019), reddit has important differences from other platforms. Unlike reddit, platforms such as Twitter and Facebook lack explicit communities with well-defined membership. reddit has a stronger focus on link-sharing than other platforms, in part due to technical differences and in part due to history and the culture on the platform. reddit is also almost entirely public, unlike private communities on Slack, Discord and some Facebook Groups.

Reviews of research conducted using reddit data have found that the generalizability of results from reddit to other contexts are promising, with certain caveats, and this is an important area for additional research (Proferes et al. 2021). However, we believe that the size and prominence of reddit means that our work still has an important impact, regardless of generalizability.

reddit also differs dramatically from peer production communities such as Wikipedia and open source projects, where the community has a clear focus beyond social engagement and entertainment. Future work is needed to understand how community values may differ in these other contexts.

reddit also mostly consists of English-speaking users from Western cultures, and additional work is needed to understand how people from other cultures may hold different values. Some evidence for cultural differences in values has already been reported (Jiang et al. 2021).

## 5 Related Work & Comparison to Existing Taxonomies

**Community Rules, Norms, & Content Moderation.** One way in which commonly-held community values manifest is in the formalized rules of communities and how those rules are enforced by content moderators. Rules on reddit have been examined in prior work (Fiesler et al. 2018; Chandrasekharan et al. 2018), however, as subreddits’ rules are written and enforced exclusively by community moderators, they may not be representative of the values of the broader community membership (Matias 2019a). On reddit, as on most social media platforms, more democratic self-moderation is relatively rare (Seering 2020), with platforms’ technical features built mostly around a strictly defined hierarchy of admins and moderators, encouraging an ‘implicit feudalism’ (Schneider 2021). In these cases, conflict between non-moderator membership and the moderators over rules and their enforcement is relatively common (Matias 2019a; Jhaver et al. 2019; Srinivasan et al. 2019). While some third party tools to enable a broader range of governance practices have been proposed (Zhang, Hugh, and Bernstein 2020), these tools are not yet widely adopted. In

this work, we directly ask all community members about their values, not just moderators.

**Intracommunity Tension and Conflict.** Conflict between non-moderators and moderators is not the only form of tension within online communities. One major challenge in communities is that of integrating new members into an existing community (Robert E Kraut et al. 2012, Ch.5). This has been studied empirically on a wide range of platforms, finding that new members mostly learn community norms from experience (Cho and Wash 2021), and that once established in a community, members are less likely to change their habits (Danescu-Niculescu-Mizil et al. 2013). As a result, periods of massive growth can result in substantial change to the community (Lin et al. 2017; Halfaker et al. 2013) and frustration amongst existing members (Kiene, Monroy-Hernández, and Hill 2016). In communities that are especially focused on topics requiring special knowledge or expertise, a related tension often occurs between those with greater knowledge and those without; this has been studied on reddit in case studies for science (Jones et al. 2019) and history communities (Gilbert 2020). In cases of extreme intracommunity conflict, such tensions can even lead to fragmentation as some members exit (Frey and Schneider 2021) and form new, alternative communities (Fiesler and Dym 2020; Newell et al. 2016). Additional evidence for intracommunity tension can be found in work that examines how members define rulebreaking and how to fairly punish such behavior; this work has found substantial disagreement amongst community members (Scheuerman et al. 2021; Jiang et al. 2021). Our taxonomy enables us to frame these tensions as conflicts between different categories of values (§4.2), or as differences in members’ value preferences in the same value category.

**Implicit Values in Prior Research.** Any research that seeks to improve the health of an online community implicitly (and often explicitly) values certain aspects of that community. For example, the abundance of research on reducing misinformation (Bovet and Makse 2019; Weld, Glenski, and Althoff 2021; Grinberg et al. 2019; Anagnostopoulos et al. 2015) implicitly values the veracity of content. Similarly, research seeking to reduce harassment (Burke Winkelman et al. 2015; Matias, Simko, and Reddan 2020) implicitly values safety. This ‘implicit values’ perspective can be used to identify what values of online communities are most studied by the research community, and conversely, what values are most understudied (§4.1).

In this work, we do not argue that the set of values derived from community members’ responses is superior to any other set of values, but instead conduct a survey to outline the diversity of values held by community members, how they may be in conflict with one another, and how they relate to prior research and can inform future research.

**Comparison to Existing Taxonomies.** While we are not aware of any work prior to ours that directly asks members of online communities their values, aspects of online interactions have been studied in the context of 1:1 or small group interactions (Deri et al. (2018, Table 1) provides an

This Work	Bao et al. (2021)	Deri et al. (2018) and Choi et al. (2020)	Fiesler et al. (2018)
Iterative categorization of 212 redditors' survey responses	Survey of Prosocial Behavior literature	Survey of Sociology & Psychology literature	Manual coding of 300 rules from 18 subreddits
Quality of Content	Information Sharing	Knowledge, Fun	Content/Behavior, Format, Low-Quality Content, Off-topic, Reposting, Spam, Advertising/Commercialization, Images
Community Engagement	Social Cohesion, Social Support, Gratitude, Mentoring, Esteem Enhancement	Power, Status, Support	Personality
Diversity	Social Cohesion	Similarity, Identity	Off-topic
Size	Social Cohesion	N/A	N/A
Participation & Inclusion	Social Support, Mentoring, Absence of Antisocial Behavior	Trust, Identity, Conflict	Doxxing/Personal Info, Harassment, Hate Speech, Trolling, Links & Outside Content
Technical Features	N/A	N/A	NSFW
Moderation & Moderators	N/A	Power, Status	Consequences/Moderation/ Enforcement
Norms	Absence of Antisocial Behavior	Trust, Conflict	Format, Voting
Trust	Information Sharing	Trust	Content/Behavior
Not mentioned by community members	Fundraising & Donating	Romance	Copyright/Piracy, Personal Army, Politics, Sitewide

Table 2: A comparison of how categories from our empirically-derived taxonomy are mapped onto by taxonomies from prior work on small group interactions (Bao et al. 2021; Deri et al. 2018; Choi et al. 2020) and community rules (Fiesler et al. 2018). All of our categories have analogues in these other taxonomies from different contexts.

overview). We summarize how the major categories from our results (§3) relate to categories in taxonomies of aspects of small group interactions (Bao et al. 2021; Deri et al. 2018; Choi et al. 2020) and the taxonomy of subreddit rules from Fiesler et al. (2018) in Table 2.

Every major category in our taxonomy has at least one analogue in the small group interaction and rule taxonomies. Naturally, given the different contexts, there are differences in how and where the categories overlap, and some components of the other taxonomies are not as relevant to the large online community context we study here. The small group taxonomies emphasize dynamics that affect 1:1 interactions, such as gratitude, mentorship, and asymmetrical power dynamics. In large communities, these translate to the quality of community engagement and interaction, the inclusion of new members, and the power wielded by moderators, respectively. Romance and Fundraising/Donating, which appear in the small group taxonomies, do not appear in our taxonomy and were not mentioned by any of our 1,481 participants' idea units.

The taxonomy of different rules on reddit (Fiesler et al. 2018) also overlaps substantially with our taxonomy, as rules are a formalized reflection of community values. As rules are primarily written with regards to content (Fiesler et al. 2018), our *Quality of Content* category is relevant to many categories from the rules taxonomy. Categories from the rules taxonomy which do not have analogues in our taxonomy are rules regarding copyright/piracy, politics, and 'personal armies' (brigading, when large numbers of members from one community temporarily participate in an often hostile manner in another community).

## 6 Conclusion

Online social communities are rich spaces that can bring people together in a healthy, productive, and enjoyable manner. Many researchers study how to make online communities 'better,' but understanding what 'better' means is a challenging problem, as there is no single set of values for online communities that can be used to inform research in this space. The values held by community members themselves are difficult to measure, and their perspectives have mostly not been included in existing research.

In this work, we surveyed 212 redditors who are members of 627 unique communities. Using open ended questions (§2.2), we asked these redditors what their values for their communities are, in their own words. Using an iterative categorization method based in grounded theory (§2.5), we contributed a taxonomy of 29 subcategories of community values across a broad range of topics from the diversity of the community to technical features (§3). Raters using our codebook demonstrate very high inter-rater agreement (Fleiss kappa = 0.874).

Our findings have important implications for future work on online social communities, and have already enabled followup work (Weld, Zhang, and Althoff 2022; Maftouni, Dubois, and Bunt 2022). We highlighted understudied and challenging-to-implement community values such as *Quality of Content*, *Size*, and *Community Engagement* (§4.1). We identified where community values conflict with one another (§4.2), and called for additional work on participatory governance for online communities that protects vulnerable groups of community members (§4.3).

## Acknowledgements

This research was supported by the Office for Naval Research (#N00014-21-1-2154), NSF grant IIS-1901386, NSF CAREER IIS-2142794, NSF grant CNS-2025022, the Bill & Melinda Gates Foundation (INV-004841), and a Microsoft AI for Accessibility grant.

## References

Allison, K.; and Bussey, K. 2020. Communal Quirks and Circle-jerks: A Taxonomy of Processes Contributing to Insularity in Online Communities. *ICWSM*.

Anagnostopoulos, A.; Bessi, A.; Caldarelli, G.; Vicario, M. D.; Petroni, F.; Scala, A.; Zollo, F.; and Quattrociocchi, W. 2015. Viral Misinformation: The Role of Homophily and Polarization. *WWW Companion*.

Bao, J.; Wu, J.; Zhang, Y.; Chandrasekharan, E.; and Jurgens, D. 2021. Conversations Gone Alright: Quantifying and Predicting Prosocial Outcomes in Online Conversations. *TheWebConf*.

Barthiel, M.; Stocking, G.; Holecomb, J.; and Mitchell, A. 2016. Reddit news users more likely to be male, young and digital in their news preferences. *Pew Research Center*.

Baumgartner, J.; Zannettou, S.; Keegan, B.; Squire, M.; and Blackburn, J. 2020. The Pushshift Reddit Dataset. *arXiv:2001.08435*.

Bovet, A.; and Makse, H. A. 2019. Influence of fake news in Twitter during the 2016 US presidential election. *Nature Communications*.

Burke Winkelman, S.; Oomen-Early, J.; Walker, A.; Chu, L.; and Yick-Flanagan, A. 2015. Exploring Cyber Harassment among Women Who Use Social Media. *Univ. J. of Public Health*.

Chandrasekharan, E.; Pavalanathan, U.; Srinivasan, A.; Glynn, A.; Eisenstein, J.; and Gilbert, E. 2017. You Can't Stay Here: The Efficacy of Reddit's 2015 Ban Examined Through Hate Speech. *Proc. ACM Hum.-Comput. Interact.*, 1(CSCW).

Chandrasekharan, E.; Samory, M.; Jhaver, S.; Charvat, H.; Bruckman, A.; Lampe, C.; Eisenstein, J.; and Gilbert, E. 2018. The Internet's Hidden Rules: An Empirical Study of Reddit Norm Violations at Micro, Meso, and Macro Scales. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW): 1–25.

Cho, J.; and Wash, R. 2021. How Potential New Members Approach an Online Community. *CSCW*.

Choi, M.; Aiello, L. M.; Varga, K. Z.; and Quercia, D. 2020. Ten Social Dimensions of Conversations and Relationships. *WWW*.

Ciampaglia, G.; and Taraborelli, D. 2015. MoodBar: Increasing New User Retention in Wikipedia through Lightweight Socialization. *CSCW*.

Danescu-Niculescu-Mizil, C.; West, R.; Jurafsky, D.; Leskovec, J.; and Potts, C. 2013. No Country for Old Members: User Lifecycle and Linguistic Change in Online Communities. *WWW*.

Deri, S.; Rappaz, J.; Aiello, L. M.; and Quercia, D. 2018. Coloring in the Links: Capturing Social Ties as They are Perceived. *CSCW*.

Erickson, T.; and Kellogg, W. 2000. Social translucence: an approach to designing systems that support social processes. *TOCHI*.

Fiesler, C.; and Dym, B. 2020. Moving Across Lands: Online Platform Migration in Fandom Communities. *CSCW*.

Fiesler, C.; Jiang, J. A.; McCann, J.; Frye, K.; and Brubaker, J. 2018. Reddit Rules! Characterizing an Ecosystem of Governance. *ICWSM*.

Fiesler, C.; and Proferes, N. 2018. "Participant" Perceptions of Twitter Research Ethics. *Social Media + Society*.

FORCE11. 2020. The FAIR Data principles.

Frey, S.; and Schneider, N. 2021. Effective Voice: Beyond Exit and Affect in Online Communities. *arXiv:2009.12470*.

Friedman, B.; Kahn, P. H.; and Borning, A. 2006. Value Sensitive Design and Information Systems. *Human-Computer Interaction and Management Information Systems*.

Gebru, T.; Morgenstern, J.; Vecchione, B.; Vaughan, J. W.; Walach, H.; Iii, H. D.; and Crawford, K. 2021. Datasheets for datasets. *CACM*.

Gilbert, E. 2012. Designing Social Translucence over Social Networks. *CHI*.

Gilbert, S. A. 2020. "I Run the World's Largest Historical Outreach Project and It's on a Cesspool of a Website." Moderating a Public Scholarship Site on Reddit: A Case Study of r/AskHistorians. *CSCW*.

Glaser, B.; Strauss, A.; and Strutzel, E. 1968. The Discovery of Grounded Theory: Strategies for Qualitative Research. *Nursing Research*.

Grinberg, N.; Joseph, K.; Friedland, L.; Swire-Thompson, B.; and Lazer, D. 2019. Fake news on Twitter during the 2016 U.S. presidential election. *Science*.

Guinier, L. 1994. *The tyranny of the majority : fundamental fairness in representative democracy*. Free Press.

Gwet, K. L. 2014. *Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Raters*. Advanced Analytics.

Habib, H.; Musa, M. B.; Zaffar, F.; and Nithyanand, R. 2019. To Act or React: Investigating Proactive Strategies For Online Community Moderation. *arXiv:1906.11932 [cs]*.

Halfaker, A.; Geiger, R. S.; Morgan, J. T.; and Riedl, J. 2013. The Rise and Decline of an Open Collaboration System: How Wikipedia's Reaction to Popularity Is Causing Its Decline. *American Behavioral Scientist*.

Halfaker, A.; Kittur, A.; and Riedl, J. 2011. Don't Bite the Newbies: How Reverts Affect the Quantity and Quality of Wikipedia Work. *WikiSym*.

Im, J.; Zhang, A. X.; Schilling, C. J.; and Karger, D. 2018. Deliberation and Resolution on Wikipedia: A Case Study of Requests for Comments. *CSCW*.

Jhaver, S.; Appling, D. S.; Gilbert, E.; and Bruckman, A. 2019. "Did You Suspect the Post Would Be Removed?": Understanding User Reactions to Content Removals on Reddit. *CSCW*.

Jhaver, S.; Bruckman, A.; and Gilbert, E. 2019. Does Transparency in Moderation Really Matter? User Behavior After Content Removal Explanations on Reddit. *CSCW*.

Jhaver, S.; Frey, S.; and Zhang, A. 2021. Designing for Multiple Centers of Power: A Taxonomy of Multi-level Governance in Online Social Platforms. *arXiv:2108.12529*.

Jhaver, S.; Ghoshal, S.; Bruckman, A.; and Gilbert, E. 2018. Online Harassment and Content Moderation: The Case of Blocklists. *TOCHI*.

Jhaver, S.; Vora, P.; and Bruckman, A. 2017. Designing for Civil Conversations: Lessons Learned from ChangeMyView. *GVU Technical Reports*.

Jiang, J. A.; Scheuerman, M.; Fiesler, C.; and Brubaker, J. 2021. Understanding international perceptions of the severity of harmful content online. *PLoS ONE*.

Jones, R.; Colusso, L.; Reinecke, K.; and Hsieh, G. 2019. R/Science: Challenges and Opportunities in Online Science Communication. *CHI*.

Kamalpour, M.; Watson, J.; and Buys, L. 2020. How Can Online Communities Support Resilience Factors among Older Adults. *Int. J. of Human-Computer Interaction*.

Kelty, C. 2017. Too Much Democracy in All the Wrong Places: Toward a Grammar of Participation. *Current Anthropology*.

Kiene, C.; and Hill, B. M. 2020. Who Uses Bots? A Statistical Analysis of Bot Usage in Moderation Teams. *CHI EA*.

Kiene, C.; Monroy-Hernández, A.; and Hill, B. M. 2016. Surviving an "Eternal September": How an Online Community Managed a Surge of Newcomers. *CHI*.

Konieczny, P. 2017. Decision making in the self-evolved collegiate court: Wikipedia's Arbitration Committee and its implications for self-governance and judiciary in cyberspace. *Intl. Sociology*.

Lakkaraju, H.; McAuley, J.; and Leskovec, J. 2013. What's in a Name? Understanding the Interplay between Titles, Content, and Communities in Social Media. *ICWSM*.

Landis, J. R.; and Koch, G. G. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics*.

Lenhart, A.; Ybarra, M.; Zickuhr, K.; and Price-Feeney, M. 2016. Online Harassment, Digital Abuse, and Cyberstalking in America. Technical report, Data & Society Research Institute.

Lin, Z.; Salehi, N.; Yao, B.; Chen, Y.; and Bernstein, M. S. 2017. Better When It Was Smaller? Community Content and Behavior After Massive Growth. *ICWSM*.

MacQueen, K.; McLellan, E.; Kay, K.; and Milstein, B. 1998. Codebook Development for Team-Based Qualitative Analysis. *Field Methods*.

Maftouni, M.; Dubois, P. M. J.; and Bunt, A. 2022. "Thank you for being nice": Investigating Perspectives Towards Social Feedback on Stack Overflow. *Graphics Interface*.

Matias, J. N. 2019a. The Civic Labor of Volunteer Moderators Online. *Social Media + Society*.

Matias, J. N. 2019b. Preventing harassment and increasing group participation through social norms in 2,190 online science discussions. *PNAS*.

Matias, J. N.; Simko, T.; and Reddan, M. 2020. Reducing the Silencing Role of Harassment in Online Feminism Discussions.

Medvedev, A. N.; Lambiotte, R.; and Delvenne, J.-C. 2019. The anatomy of Reddit: An overview of academic research. *arXiv:1810.10881 [cs]*.

Newell, E.; Jurgens, D.; Saleem, H. M.; Vala, H.; Sassine, J.; Armstrong, C.; and Ruths, D. 2016. User Migration in Online Social Networks: A Case Study on Reddit During a Period of Community Unrest. *ICWSM*.

Proferes, N.; Jones, N.; Gilbert, S.; Fiesler, C.; and Zimmer, M. 2021. Studying Reddit: A Systematic Overview of Disciplines, Approaches, Methods, and Ethics. *Social Media + Society*.

Raj, E. X.; and Daniels, D. E. 2017. Psychosocial support for adults who stutter: Exploring the role of online communities. *Speech, Language and Hearing*.

Robert E Kraut; Paul Resnick; Sara Kiesler; Moira Burke; Yan Chen; Niki Kittur; Joseph Konstan; Yuqing Ren; and John Riedl. 2012. *Building Successful Online Communities: Evidence-Based Social Design*. The MIT Press.

Santos, T.; Burghardt, K.; Lerman, K.; and Helic, D. 2020. Can Badges Foster a More Welcoming Culture on Q&A Boards? *ICWSM*.

Scheuerman, M.; Jiang, J. A.; Fiesler, C.; and Brubaker, J. 2021. A Framework of Severity for Harmful Content Online. *CSCW*.

Schneider, N. 2021. Admins, mods, and benevolent dictators for life: The implicit feudalism of online communities. *New Media & Society*.

Seering, J. 2020. Reconsidering Self-Moderation: The Role of Research in Supporting Community-Based Models for Online Content Moderation. *CSCW*.

Shen, Q.; and Rose, C. 2019. The Discourse of Online Content Moderation: Investigating Polarized User Responses to Changes in Reddit's Quarantine Policy. *Proc. of the Third Workshop on Abusive Language Online*.

Srinivasan, K. B.; Danescu-Niculescu-Mizil, C.; Lee, L.; and Tan, C. 2019. Content Removal as a Moderation Strategy: Compliance and Other Outcomes in the ChangeMyView Community. *CSCW*.

Steinmacher, I.; Conte, T.; Gerosa, M. A.; and Redmiles, D. 2015. Social Barriers Faced by Newcomers Placing Their First Contribution in Open Source Software Projects. *CSCW*.

Strauss, A. L. 1987. *Qualitative analysis for social scientists*. Cambridge University Press.

Trippodi, F. 2021. Ms. Categorized: Gender, notability, and inequality on Wikipedia. *New Media & Society*.

Weld, G.; Glenski, M.; and Althoff, T. 2021. Political Bias and Factualness in News Sharing across more than 100,000 Online Communities. *ICWSM*.

Weld, G. C.; Zhang, A. X.; and Althoff, T. 2022. What Makes Online Communities 'Better'? Measuring Values, Consensus, and Conflict across Thousands of Subreddits. *ICWSM*.

Wikipedia. 2021. Wikipedia:Polling is not a substitute for discussion.

Williamson, D. A. 2020. US Social Media Usage: How the Coronavirus Is Changing Consumer Behavior. *Insider Intelligence*.

Yazdanian, R.; Zia, L.; Morgan, J.; Mansurov, B.; and West, R. 2019. Eliciting New Wikipedia Users' Interests via Automatically Mined Questionnaires: For a Warm Welcome, Not a Cold Start. *ICWSM*.

Zhang, A. X.; Hugh, G.; and Bernstein, M. S. 2020. PolicyKit: Building Governance in Online Communities. *UIST*.

Zhang, J.; Danescu-Niculescu-Mizil, C.; Dixon, L.; Hua, Y.; Thain, N.; and Taraborelli, D. 2018a. Conversations Gone Awry: Detecting Early Signs of Conversational Failure. *ACL*.

Zhang, J.; Danescu-Niculescu-Mizil, C.; Sauper, C.; and Taylor, S. J. 2018b. Characterizing Online Public Discussions through Patterns of Participant Interactions. *CSCW*.

## 7 Required Paper Checklist

### 1. For most authors...

- Would answering this research question advance science without violating social contracts, such as violating privacy norms, perpetuating unfair profiling, exacerbating the socio-economic divide, or implying disrespect to societies or cultures? **Impacts from the research are discussed in §1 and 4, ethical and privacy concerns are discussed in §2.7.**
- Do your main claims in the abstract and introduction accurately reflect the paper's contributions and scope? **Yes.**
- Do you clarify how the proposed methodological approach is appropriate for the claims made? **Yes, discussed in §2.**

(d) Do you clarify what are possible artifacts in the data used, given population-specific distributions? [Yes, discussed in §4.4.](#)

(e) Did you describe the limitations of your work? [Yes, see §4.4.](#)

(f) Did you discuss any potential negative societal impacts of your work? [Yes, discussed in §4.4.](#)

(g) Did you discuss any potential misuse of your work? [Yes, discussed in §2.7.](#)

(h) Did you describe steps taken to prevent or mitigate potential negative outcomes of the research, such as data and model documentation, data anonymization, responsible release, access control, and the reproducibility of findings? [Yes, described in §2.7.](#)

(i) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes.](#)

2. Additionally, if your study involves hypotheses testing...

- (a) Did you clearly state the assumptions underlying all theoretical results? Not applicable.
- (b) Have you provided justifications for all theoretical results? Not applicable.
- (c) Did you discuss competing hypotheses or theories that might challenge or complement your theoretical results? Not applicable.
- (d) Have you considered alternative mechanisms or explanations that might account for the same outcomes observed in your study? Not applicable.
- (e) Did you address potential biases or limitations in your theoretical framework? [Yes, addressed in 4.4.](#)
- (f) Have you related your theoretical results to the existing literature in social science? [Yes, see §5.](#)
- (g) Did you discuss the implications of your theoretical results for policy, practice, or further research in the social science domain? [Yes, discussed in §6.](#)

3. Additionally, if you are including theoretical proofs...

- (a) Did you state the full set of assumptions of all theoretical results? Not applicable.
- (b) Did you include complete proofs of all theoretical results? Not applicable.

4. Additionally, if you ran machine learning experiments...

- (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? Not applicable.
- (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? Not applicable.
- (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? Not applicable.
- (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? Not applicable.

(e) Do you justify how the proposed evaluation is sufficient and appropriate to the claims made? Not applicable.

(f) Do you discuss what is “the cost” of misclassification and fault (in)tolerance? Not applicable.

5. Additionally, if you are using existing assets (e.g., code, data, models) or curating/releasing new assets, **without compromising anonymity**...

- (a) If your work uses existing assets, did you cite the creators? [Yes, cited in §5 and 2.](#)
- (b) Did you mention the license of the assets? Not applicable.
- (c) Did you include any new assets in the supplemental material or as a URL? Not applicable.
- (d) Did you discuss whether and how consent was obtained from people whose data you’re using/curating? [Yes, discussed in §2.7.](#)
- (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [Yes, in §2.7.](#)
- (f) If you are curating or releasing new datasets, did you discuss how you intend to make your datasets FAIR (see FORCE11 (2020))? Not applicable.
- (g) If you are curating or releasing new datasets, did you create a Datasheet for the Dataset (see Gebru et al. (2021))? Not applicable.

6. Additionally, if you used crowdsourcing or conducted research with human subjects, **without compromising anonymity**...

- (a) Did you include the full text of instructions given to participants and screenshots? Not applicable.
- (b) Did you describe any potential participant risks, with mentions of Institutional Review Board (IRB) approvals? Not applicable.
- (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? Not applicable.
- (d) Did you discuss how data is stored, shared, and de-identified? Not applicable.

## A Additional Figures Describing Survey Interface and Recruiting

We retrieved 5 subreddits you participated in recently. Please remove any you don't consider yourself a member of, and edit or add any that we missed.

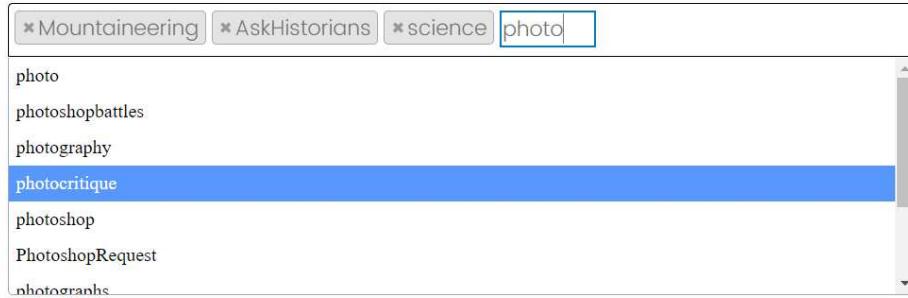


Figure 1: A screenshot of the interface used by participants to enter the subreddits they consider themselves a member of. This search box queries the reddit API in real-time to populate the results and ensure that only valid subreddit names are entered.



Figure 2: Reddit advertisements used to recruit participants.

## B Additional Recruiting and Incentive Details

Participants for this study were recruited primarily through the purchase of reddit advertisements. These advertisements display inline with other content in both individual subreddits and aggregated content views, and are shown on the reddit website as well as the official reddit mobile app. Figure 2 shows the appearance of these advertisements.

Upon clicking on the advertisement, the user is taken to the first page of the survey which summarizes the aims of the study and requests informed consent to continue. Over the course of the study, 920,025 advertisement impressions were made, generating 2,084 clicks, for a click-through rate of 0.227%. Of these people who clicked through to the first page of the survey, 509 started to complete the survey, and 212 completed it. Eleven respondents did not consent to continuing the survey, and as a result were not shown any additional questions. To increase the diversity of recruiting, the survey was also distributed to relevant university mailing lists and Slack channels at two large American universities, as well as posted to /r/SampleSize, a subreddit for the distribution of surveys, recruiting an additional 41 participants.

To incentivize respondents to participate, a raffle was held, and winners were chosen at random from the pool of respondents who both completed the survey and supplied their reddit username, which was used to contact the winners. One ‘first place’ prize, a \$100 Amazon gift card, and five ‘second place’ prizes of \$20 Amazon gift cards were awarded (equivalent amounts in local currency were provided to participants outside of the United States).

## C Additional Details on Participant Demographics

	Our Survey (n = 212)	reddit Overall Barthiel et al. (2016)
<b>Age</b>		
18-29	101 (71.1%)	59%
30-49	35 (24.6%)	33%
50-64	6 (4.2%)	7%
65 or older	0 (0.0%)	<1%
<b>Gender</b>		
Woman	49 (26.2%)	33%
Man	129 (69.0%)	67%
Non-binary	9 (4.8%)	
Additional Gender	5 (2.7%)	Not reported
<b>Race and Ethnicity</b>		
White (Hispanic, Latino, or Spanish)	32 (16.9%)	7% Hispanic
Non-white Hispanic, Latino, or Spanish	6 (3.2%)	
White (Not Hispanic, Latino, or Spanish)	89 (47.1%)	74% White
Black or African American	3 (1.6%)	8% Black
Middle Eastern or North African	8 (4.2%)	
Asian	46 (24.3%)	
Native Hawaiian or Pacific Islander	3 (1.6%)	10% Other
Indigenous American or Alaskan Native	1 (0.5%)	
Additional Race/Ethnicity	10 (5.3%)	

Table 3: Demographics of survey respondents and overall reddit demographics. In general, our respondents’ demographics are similar to the overall demographics of reddit. Participants could choose multiple gender and race/ethnicity options, so percentages may not sum to 100%. Overall reddit demographic data (Barthiel et al. 2016) uses different racial identity questions, so while an exact comparison is not possible, similar categories are provided here.

## D Survey Instrument

### Informed Consent

This survey aims to learn more about the subreddits that reddit users participate in, and what values redditors have for those subreddits.

In this survey, we will ask you a few questions about your overall reddit usage, and then ask you a few questions each about the subreddits you consider yourself to be a member of. You may skip any questions you'd prefer not to answer. It should take less than 5 minutes to complete.

Only high-level data will be published as part of our research. Your responses are confidential, and will never be made public. This study is run by researchers from the University of Washington, and has been determined to be exempt from IRB approval under University of Washington IRB ID STUDY00011457. For questions or concerns, please contact gweld@cs.washington.edu, or /u/cyclistNerd on reddit.

Would you like to participate in this survey?

- Yes, I would like to participate in this survey.
- No, I would not like to participate in this survey.

### Reddit Username and Compensation

As compensation for your participation in this study, you will be entered in a raffle for one of five Amazon gift cards, worth \$20 each.

To contact you after the raffle drawing, we will send you a reddit private message. To do so, we ask for your reddit username. We will also use your username to identify your posts and comments in the subreddits you participate in. Your username will be kept confidential, and we will never publish any of your reddit history.

Providing your username is entirely optional, but without it, we cannot enter you into the raffle.

What is your primary reddit username? Your answer will be kept confidential.

Please spell carefully, and do not include '/u/' or 'u/'.

Do you have multiple reddit accounts?

- No, I only have one reddit account.
- Yes, I have multiple reddit accounts.

You entered that your primary username is /u/

If this looks correct, press next to start the rest of the survey. If this is incorrect, please press back to edit your response.

### Multiple Accounts

Earlier, you indicated that you had multiple reddit accounts.

If you're comfortable doing so, please enter the names of your alternate reddit accounts, separated by commas.

### Demographics

What is your age (in years)?

Please describe your gender (check all that apply)

- Woman
- Man
- Non-binary
- Prefer to self-describe

Please describe your race (check all that apply)

- White (Hispanic, Latino, or Spanish)
- White (Not Hispanic, Latino, or Spanish)
- Non-white Hispanic, Latino, or Spanish
- Black or African American
- Asian
- Middle Eastern or North African
- Native Hawaiian or Pacific Islander
- Indigenous American or Alaskan Native
- Prefer to self-describe

## Overall Reddit Usage

In this section, we will ask you about how you use reddit.

Typically, how often do you use reddit?

- Every day.
- A few times a week.
- Once a week.

Typically, how long do you typically spend on reddit at one time?

- Less than five minutes.
- More than five minutes, but less than fifteen minutes.
- More than fifteen minutes, but less than an hour.
- More than an hour.

Typically, how often do you post or comment on reddit versus browsing what others have submitted (lurking)?

- Frequently, I frequently submit posts or comment on threads.
- Occasionally, I post or comment occasionally, but mostly browse what others have submitted.
- Rarely, I almost always just browse what others have submitted.
- Never, I only browse what others have submitted.

How often do you use aggregate subreddits (like your frontpage, /r/all, or multi-reddits) versus looking at individual subreddits?

- Always, I never look at individual subreddits.
- Frequently, I mostly use aggregate subreddits, but sometimes I look at individual subreddits.
- Occasionally, I look at aggregate subreddits and individual subreddits about evenly.
- Rarely, I mostly look at individual subreddits.
- Never, I only look at individual subreddits.

Do you use reddit more from your computer or from your phone?

- I only use reddit from my computer.
- I mostly use my computer, but sometimes use my phone.
- I use my phone and my computer about evenly.
- I mostly use my phone, but sometimes use my computer.
- I only use reddit from my phone.

## Subreddit Selection

In this section, we'll ask you questions specific to the subreddits that you consider yourself a member of. These subreddits should be subreddits that are important to you, and that you are familiar enough with to feel comfortable commenting on different aspects of how they are run, and how their members interact with one another.

Thank you for selecting subreddits. The next section will ask you about your experiences with each subreddit, individually.

**Open Ended Subreddit Value Questions** In this section, we'll ask you questions specifically about your experience in /r/\_\_\_\_\_.

In your responses, please consider not only the content of /r/\_\_\_\_\_, but also how it is run, how its members treat one another, and anything else that impacts your experience.

As it exists right now, what are a few of the best aspects of the /r/\_\_\_\_\_ community?

If you could change anything, what are some aspects of the /r/\_\_\_\_\_ community you would like to improve upon?

### Reflection Questions

These last two questions ask you to reflect on your experience across all the subreddits you've used.

Generally, what values do you think are important in an online community? What makes a community healthy?

Have you ever stopped participating in any subreddits? What are some signs that a community isn't worth your time, or isn't a community you want to participate in?

## E Codebook

### Introduction

This taxonomy is intended to classify individual idea units, which are focused on a specific aspect of a community. As such, it is intended to be mutually exclusive - idea units should not be assigned to multiple categories. If a participants' response appears to belong to multiple categories, it likely should be subdivided into multiple idea units.

Furthermore, the taxonomy is hierarchical, with the categories being grouped into 9 high level groups. Idea units should be tagged with the most specific category possible - while the high level group is also available as a label itself, it should only be applied when the idea unit in question is either a) too short/vague to be assigned to a more specific category, or b) does not clearly fit into any of the subcategories.

### Taxonomy Categories and Descriptions

**1) Quality of Content** Idea units that belong to this category should be commenting on the perceived value/utility of the content in the subreddit, or lack thereof, with exceptions for idea units regarding bullying/offensive content (these belong in category 5a) or idea units regarding the trustworthiness of the content (these belong in category 9b). As with all categories, idea units in the Quality of Content category should be assigned to as specific of a category as possible. Thus, many of the idea units that fall into category 1 (as opposed to subcategories 1a-d) are fairly vague/generic.

#### Examples

*I like the content*

*Content is easy to understand*

#### Counterexamples

*The content is funny* belongs in 1a due to the specificity of “funny” (it’s entertaining).

*There are not enough memes* belongs in 1b because the desire for memes is a personal preference.

*Some posts are degrading to women* belongs in 5a because it refers specifically to content this offensive or abusive.

**1a) Education, Entertainment** This category should contain idea units that comment on the value of the content, be it educational, entertaining, or some combination (such as “interesting”). It can sometimes be difficult to distinguish idea units in this category from those in 1b Personal Preferences. You should use your judgment to determine if the survey respondent’s emphasis is more on the utility of the content (belonging in 1a) the the specific type of content they prefer (belonging in 1b).

#### Examples

*The posts are funny*

*I learn a lot about different mushrooms*

#### Counterexamples

*I especially laugh at the pictures of silly cats* belongs in 1b because the emphasis is on the personal preference towards a specific type of content (silly cat pictures).

**1b) Personal Preferences** This category should contain idea units that reflect the respondent’s preferences towards (or against) particular types of content, frequently including idea units on memes, specific subject matter, etc.

#### Examples

*Good memes*

*Custom gaming PCs*

*Less discussion about politics would be good*

**1c) Curation, Recency, Discovery** This category focuses on how the content in the subreddit helps the community member discover new things that they may not have otherwise seen, or helps them make sense of large volumes of content (curation).

#### Examples

*/r/seattle helps me keep up to date with news*

*/r/hiphopheads helps me find the best new music*

**1d) Spam, Reposts, Bots** Idea units in this category focus on the presence of specific types of content such as spam, reposts, and content submitted by bots.

#### Examples

*There are too many reposts in /r/pics*

*I don't like the repetitive questions*

*The community bot is helpful*

*Too much spam*

## 2) Community Engagement

Idea units in this category focus not on the content, but on the community, either the community as a whole abstract entity, or on specific people or groups of people. One exception to this is idea units about the credentials or knowledge of people in the community, which belong in 9a Knowledgeable People. Most idea units in this category are likely to fall into the more specific idea units should be assigned to subcategories 2a and 2b, while very generic or high level idea units should be assigned to this top level category.

Examples

*Community*

Counterexamples

*Friendly community* belongs in 2a as it comments on a quality (friendliness) of the community as a whole.

**2a) Quality of Interaction or Community as a Whole** Idea units in this category comment on both good and bad aspects of the community or of specific interactions within the community. Note that some specific types of interactions belong in other categories (e.g. size in 4b, harassment in 5a, voting behavior in 8a, rule/norm-breaking in 8b).

Examples

*My posts are replied to quickly*

*The community is nice*

*I get to discuss my favorite TV show [in /r/TwinPeaks]*

**2b) Connection, Universalization** Idea units in this category should focus on the impact of the community on the survey respondent, such as feeling connected to community members, or aware that there are others out there who feel the same as they do (universalization).

Examples

*/r/meow\_irl is so relatable*

*/r/mentalhealth makes me feel connected to others*

*/r/raisedbynarcissists makes me know there are others who feel like me*

**3) Diversity** This category focuses on the diversity of content and people within a community. Almost all idea units in this category should be able to be assigned to one of the two subcategories 3a and 3b.

**3a) Variety of Content** This category focuses on the variety of content. All idea units which praise or critique the diversity of content should fall within this category, with the exception of complaints about reposts (re-submission of the exact same content) which belong in 1d.

Examples

*I get to see so many different kinds of snails [on /r/snails]*

*I don't like how there is a hivemind*

**3b) Diversity of People** This category should contain idea units on the diversity of the people within the community or the community as a whole. Idea units that focus on the diversity (or lack thereof) peoples' opinions or ideas also belong in this category.

Examples

*There are people of so many different backgrounds*

*I don't like the hivemind about some musicians*

**4) Size** Idea units in the size category will most likely fall into either of the two subcategories unless they are very vague.

Counterexamples

*Too big* should be assigned to 4b as adjectives “big” and “small” generally refer to the size of the community unless explicitly stated otherwise.

**4a) Volume of Content** This category contains idea units that relate to the volume of content (including posts and comments) within a subreddit, or the rate at which this content is posted.

Examples

*I wish there were more posts*

*Not enough people post*

*I like that there are always new posts for me to look at*

*Posts get lots of comments*

*Posts (or comments) are submitted so frequently*

**4b) Size of Community** This category contains idea units that relate to the size of the community (i.e. the number of people who are in the community or who participate in the community)

Examples

*I wish the community were bigger*

*The community is small and close-knit* belongs in 4b as the emphasis is primarily on the size of the community. Idea units in the category can be similar to those in 2a Quality of Interaction or Community as a Whole. Use your judgment to determine if the primary emphasis is on the size of the community or on some other qualities

Counterexamples

*The community is close-knit* belongs in 2a as the primary emphasis of the idea unit is on the connection within the community, not the size.

**5) Participation and Inclusion** Idea units in this category should focus on who participates in the community, and actions and content that explicitly impact who participates and who is included. Idea units that mention aspects of a community that may have an impact on inclusion, but this impact is not explicitly stated (e.g. ‘friendly people’) should be categorized in 2a.

Examples

*Everyone is included*

Counterexamples

*They are mean to new members* belongs in subcategory 5a as it relates to behavior which is specifically abusive or harassing

**5a) Offensive, Abusive, Harassing Content or Behaviors** Idea units in this category relate to content and behaviors that are offensive, abusive, or harassing to specific people, groups of people, or in general. You should use your best judgment to interpret if the content is offensive to the respondent, in which case it belongs in this category, or if it is simply not in alignment with their personal preferences, in which case it belongs in 1b.

Examples

*I wish there were less jokes that make fun of women* is offensive

*Bullying*

*[/r/pics] pictures of male genitalia* belongs in this category because in the context of /r/pics, pictures of genitalia are unexpected and therefore likely offensive.

Counterexamples

*[/r/nudes] pictures of male genitalia* does not belong in this category because the most likely interpretation of “pictures of male genitalia” in the context of /r/nudes is that of a personal preference (given that such pictures would be expected on the /r/nudes subreddit) and therefore this idea unit is best categorized under 1b.

*People are quick to criticize* belongs in 5 because criticism isn’t inherently offensive

**5b) Outsiders and Demographics** Idea units in this category should explicitly relate to who participates in the community, and their perceived out/in-group status.

Examples

*[/r/wichita] too many people from outside the city post here*

*Posts are mostly submitted by people from outside the community*

**5c) Tools for Participation** This category relates to tools (such as wikis, FAQs, and stickied posts) that help improve participation.

Examples

*The bot that welcomes new members is really nice*

*There should be an FAQ explaining how to post*

**6) Technical Features** Idea units in this category and subcategories should relate to technical features offered (or not offered by reddit). As with all categories, the most specific subcategory should be used. As such, this top-level category should only be used for broad idea units that do not fit into 6a-c.

Examples

*Videos don’t load sometimes*

*I prefer old reddit to the redesign*

*The mobile app works well*

**6a) Flairs, Tags, NSFW Labels** Idea units in this category should relate to flairs (small icons or text strings associated with usernames), tags applied to posts (this includes stickied/pinned posts or comments), and NSFW (Not Safe for Work) labels.

Examples

*[/r/colonoscopy] NSFW tags are really well used*

*Flairs are helpful for knowing who is who*

*Post tags for categories are convenient*

*Sometimes NSFW content isn’t marked as such*

**6b) Search, Filters** Idea units in this category relate to searching for content or filtering for specific types of content.

Examples

*I like that I can easily search for pictures*

*I wish I could filter by experience level*

*It would be nice to be hide NSFW content* belongs in this category because it relates primarily to the ability to filter out NSFW content, not the NSFW label itself.

**6c) Recommendation Systems** Idea units in this category relate to recommendation systems for finding similar content.

Examples

*I wish there were recommendations for seeing more posts like the ones I like*

**7) Moderation and Moderators** This category contains idea units regarding who the moderators are, and how they perform their job (or fail to perform their job).

On reddit, moderators are responsible for:

- Setting rules
- Enforcing rules, including
- Removing posts
- Banning users
- Updating tags and flairs and NSFW labels
- Recruiting and selecting new moderators
- Communicating their work to the community

Note that this category should only contain idea units relating specifically to moderators and their actions. Idea units regarding rules themselves (e.g. requests for new rules) should go in 8 and its subcategories.

Examples

*Moderators are very friendly and reply quickly to messages*

*The mods are corrupt and on a power-trip*

*I wish the community was involved in new rules being created*

*Rules aren't well enforced*

*Moderators enforce rules unfairly*

Counterexamples

*I wish content from 4chan was banned* belongs in 8b because it is a request for a rule.

**8) Norms and Rules** Idea units in this category and its subcategories should relate to rules and norms, explicitly stated or not. Idea units relating to how rules are enforced, or other aspects of moderation, belong in 7 Moderators.

**8a) Voting Behavior** Idea units in this category should relate specifically and explicitly to voting behavior (upvotes and downvotes on reddit).

Examples

*People shouldn't use downvote as a disagree button*

Counterexamples

*Upvotes help me find the best music* belongs in 1c because it is primarily relating the the curation mechanism provided by votes, not how people vote.

**8b) Adherence to Norms and Rules** Idea units in this category should relate to the subreddits' rules and norms and how the community adheres to norms and rules. Many idea units in this category will be specific to the type of content within the subreddit, so checking the subreddit the idea unit corresponds to may be helpful. This category should also include requests for rules. Implicit norms can be particularly complex and context dependent, so consider these carefully and use your best judgment. Note that idea units relating to the enforcement of rules belong in 7 Moderators.

Examples

*It is frustrating when newbs ask questions that are already answered in the FAQ*

*[/r/lgbtq] People sometimes ask us to basically pick a label for them - this is a deeply personal identity, not something we can pick for other people*

*I wish content from 9gag was banned*

*They should allow memes*

*[/r/birdspotting] People should be required to include the location the saw the bird in their request for ID*

*[/r/ElderScrolls] People should specify which version of the game they're referring to*

Counterexamples

*Rules aren't enforced strictly enough* belongs in 7 because it primarily relates to enforcement, which is performed by moderators.

**9) Trust** Idea units in this category relate to the trustworthiness (or lack thereof) of the people and the content within subreddits. The top-level category should be used only when it is not possible to differentiate between idea units commenting on the people (9a) and the content (9b).

Examples

*Students post their honest experiences and opinions so it's legit* belongs in the top-level category because it is commenting both on the people (students are honest) and the content (posts are honest)

**9a) Knowledgeable People** Idea units in this category should relate to the knowledge, trustworthiness, or credentials of the people in the community.

Examples

*There are lots of doctors so I know what they're saying is correct*

*People are mostly honest*

**9b) Trustworthy Content** Idea units in this category should relate to the veracity or trustworthiness of content in the community.

Examples

*There is no fake news*

*[/r/whatisthispant] Multiple people often confirm the same ID so I know it's right*

## 10) Exclude/Unintelligible/Off-topic

### Notes for Raters

Responses are confidential. Please do not disclose them. Use your best judgment, and refer back to this codebook as necessary. Feel free to use the comments column to note any responses you were especially uncertain of. When labeling, mark the number and letter of the category in the label column. For example, an idea unit that you classify into “Tools for Participation” should be labeled “5c”. Don’t worry if the labels are deeply skewed - in our initial sample, categories 1a, 1b, and 2a were by far the most common. Leave blank idea units that you think are not categorizable, and add a comment explaining why.