No Community Can Do Everything: Why People Participate in Similar Online Communities

ANONYMOUS AUTHOR(S)
SUBMISSION ID: 2272*

Large-scale quantitative analyses have shown that individuals frequently talk to each other about similar things in different online spaces. Why do these overlapping communities exist? We provide an answer grounded in the analysis of 20 interviews with active participants in clusters of highly related subreddits: Within a broad topical area, there are a diversity of benefits that an online community can confer. These include: (a) specific information and discussions, (b) socialization with similar others, and (c) attention from the largest possible audience. A single community cannot meet every need. Our findings suggest that growing areas of interest on an platform for online communities tend to become populated by groups of specialized communities with diverse membership sizes, topical boundaries, and rules. Such systems of overlapping communities are able to provide a greater range of benefits than any single community.

CCS Concepts: • **Computer systems organization** → **Embedded systems**; *Redundancy*; Robotics; • **Networks** → Network reliability.

Additional Key Words and Phrases: datasets, neural networks, gaze detection, text tagging

1 INTRODUCTION

Early work in social computing treated online communities as isolated units whose internal processes and members' experiences could be understood without considering their participation in other online communities. Now, as community hosting platforms like Reddit and Facebook have grown in prominence, social computing scholars have sought to document and explore the connections between online communities [17, 32, 68, 80]. This research has shown that online communities

overlap with each other enormously in terms of their memberships and topics in ways that have important consequences for a range of outcomes [13, 69, 73] .

User and topic overlap is widespread—both within platforms and across them. For example, a range of studies have highlighted the fact that members frequently participate in multiple online communities. This occurs both serially as users migrate between communities over time [49, 67, 68] as well as in parallel as individuals participate in multiple groups simultaneously [35, 73, 81]. Many large platforms host distinct communities with similar topics and content [18, 80]. In at least one study, researchers have documented that overlaps in users and topics often coincide [18]. In other words, members of online communities often simultaneously participate in overlapping conversations with overlapping groups of people in different online groups.

Why are the same individuals talking to each other about similar things in different online spaces? For social computing scholars, this question reflects an empirical puzzle and points to an important gap in social computing theory. Although social computing offers many theories of why individuals might want to participate in a community, almost all empirical work in social computing on user and topic overlap has been supported by computational and quantitative analysis. As a result, we know very little about how overlaps feel to participants or what quantitative measures of overlaps mean to users. Critically, we also have very little in the way of empirical evidence that is able to speak to why communities overlap in userbases or topics in the first place.

Our work seeks to complement existing quantitative research on intercommunity relationships with qualitative richness and depth. A better qualitative understanding of intercommunity overlap contributes to several streams of social computing scholarship related to relationships between different online communities. In particular, our work complements a series of social computing studies that have taken inspiration from ecological theory and shown that online groups' growth and survival are closely tied to activity in adjacent online spaces [69, 73, 80, 81].

We seek to answer the research question in italics above through an interview-based study of Reddit users with experience in clusters of overlapping communities. Using a dataset of posts and comments on Reddit, we identify groups of communities on Reddit with highly overlapping users and topics and recruit a set of 20 participants from 9 clusters covering a variety of interest-based topics. Drawing from a grounded theory analysis of interview transcripts, develop a grounded theory that explains why a user may want to simultaneously participate in communities with overlapping userbases and topics.

Our findings suggest that users seek three salient of benefits from online groups: users want to (a) find specific types of content, discussions, and information; (b) connect with similar types of people; and (c) share content with the largest possible audience. Our work also suggests that these three things are frequently in conflict such that the more a community provides one, the less able it may be to provide the other two. Because it is difficult or impossible for a single community to provide all three benefits, these clusters of multiple overlapping communities exist to do so in aggregate.

Our work makes both both empirical and theoretical contributions to a mostly quantitative body of scholarship on intercommunity relationships by providing a qualitative sense of what relationships between online communities involve in socially meaningful and emic terms. In this sense, our work describes *how* communities relate to each other. Additionally, our work makes a theoretical contribution by describing a set of reasons *why* individuals might choose to participate in multiple online communities with similar users or topics as well as *why* platforms might want to host communities with similar, or even identical, topics.

2 RELATED WORK

2.1 Overlapping online communities

Although most research in online communities analyzes the internal factors driving online community success [39], a growing literature studies communities related by overlaps in topic or membership [18, 68, 69, 81]. This work has found that simultaneous engagement in multiple communities is common on large platforms that host online communities [68]. With several exceptions [e.g., 22, 35, 37, 79], this work typically takes the fact that communities overlap for granted and focuses on the consequences of overlap on outcomes like the emergence and growth of communities [8, 81] and the diffusion of types of language such as hate speech [13]. However, none of this work provides into how communities come to overlap and why these overlaps persist.

For example, researchers have investigated inter-community conflict and found that conflict between is only initiated by a very small percentage of online communities [40]. Other work has shown that communities with overlapping content

interact with each other to shift topical boundaries. Although these definitions of what content should or should not go in an online community might seem crisp and discrete, content that is cross-posted to multiple communities shapes the ongoing renegotiation of what the boundaries mean in practice [8]. Zhang et al. [77] have shown that topical boundaries between communities can also shift as similar online communities attract different kinds of users with different interests. In a similar way, Massanari [52] has argued that communities with overlapping users can influence each other's cultures. As a result, banning problematic communities from a platform like Reddit can reduce toxicity in related communities [12]. Related groups on such platforms often use meaningful prefixes and affixes to signify the role of a community compared to its relatives (e.g., r/askscience is a question answering community auxiliary to r/science) [31].

A number of studies on overlapping communities draw upon ecological theory [69, 73, 80, 81]. Ecological approaches in social computing theorize that overlaps between users and topics relate to competitive or mutualistic forces between communities and drive outcomes like growth and survival. For example, Wang et al. [73] found that sharing similar members reduced the rate of growth in Usenet groups and Zhu et al. [80] found that participation rates often suffered if there was too little or too much overlap in topic with others communities. Conversely, Zhu et al. [81] found that new wiki communities often benefit by sharing members with older communities and that high levels of user overlap were positively associated with community survival.

Although these studies have shown much about how communities relate to each other through statistical analysis of relationships among online communities, they

are naturally unable to speak to *how* participants understand the relationships between similar online communities or *why* they participate in multiple overlapping communities. The exclusively quantitative nature of ecological accounts of intercommunity relationships means that a range of potential explanations are possible.

Although we know of no qualitative examination focused directly on understanding why many overlapping communities exist, there are a series of qualitative papers that illuminate this question. Kiene et al.'s [37] interview study of moderators on overlapping subreddit and Discord communities describes how "technological frames" that community organizers learned on Reddit informed their use of bots to reproduce affordances for moderation on Discord. Fiesler and Dym [22] describe the history of online fanfiction writing communities migrating across platforms in pursuit of hospitable infrastructure. Similarly, Zhao et al. [79] describe how individuals use multiple social media platforms to meet their varied nuanced communication needs. Although their study is primarily quantitative, Zhu et al. [80] include quotes from interviews to support the emic validity of notions of competition and mutualism between groups in enterprise social media. Finally, Hwang and Foote's [35] paper seeks to explain why individuals participate in persistently small online communities on Reddit and ends with a reflection that many small communities are sustainable only because they are "nested" within other larger online communities. All told, these findings suggest a rich social process by which participants in online communities purposfully construct and move between overlapping spaces bringing their culture, experiences, and ideas with them.

The goal of our study is to build on this previous work to explain this process in detail. Although social computing scholarship has pointed to differences in affordances between platforms and ways in which users shift and migrate between communities over time, the lack of qualitative evidence drawn from participants in overlapping contexts in the same platform at the same time means that we lack a strong sense of why users choose to participate in multiple communities simultaneously. Although ecological studies discuss and attempt to quantify competition or mutualism, we know little about how participants in communities understand the relationships between their communities or if these concepts have any emic validity among participants. Our work aims to place applications of ecological theory to online commmunities on firmer ground by using qualitative methods to inductively answer these questions in terms of individual preferences, perceived benefits, and understandings of the relationships between communities.

2.2 Reasons for joining online communities

Decades of social computing research to support online communities has been driven to understand the reasons people have for belonging to online communities in terms of intrinsic and extrinsic rewards [39, 75]. It has long been recognized that different people have different motivations and that individuals often have multiple motivations themselves [71]. Motivations to join online communities include the social, informational, and material benefits they receive through their participation [7]. In terms of uses and gratifications theory, "users actively seek particular media with the goal of gratifying an existing need" [43]. Past research has shown that people seek online communities for such wide ranging reasons as

to collaborate on projects [60], recieve social support [45], cooperate with friends [71], and especially to exchange information [44, 46, 57, 61].

Other research focuses on the growth and decline of membership in online communities and surfaces motivations for why people choose not to participate [16]. Brandtzæg and Heim [6] found that a lack of trust or low quality content in the community can lead to decline in membership. Halfaker et al. [30] found that changes made to Wikipedia to manage article quality coincided with a decrease in retaining newcomers, crippling the community's growth. In general, online communities may have lifecycles as leaders of established communities become resistent to change and unwelcoming to newcomers and thereby decline [64, 70]. While these studies focus on the individual motivations for joining a single online community, our work asks how and why individuals join multiple, highly related online communities.

While our findings are the result of an inductive process of bottom-up thematic analysis from the interview data we collected, the way we present our findings relies on three theoretical concepts that each describe one of the reasons that individuals will join online communities. The following concepts are "sensatizing" concepts in our analysis as described in §4.

2.2.1 Finding specific content. One of the most important benefits of online communities is to foster the spread of useful content like knowledge and information [21]. By connecting individuals with diverse information and skills, online communities provide opportunities to match knowledge seekers with experts and to foster collaboration on information goods such as encyclopedias and software

[4, 24, 27, 41]. Indeed, research has long focused on the ways that individuals utilize diverse types of social computing systems to meet information needs through systems like Q&A sites [2], synchronous chat systems [74], search engines [55], social network sites [56, 65], fanworks [24], and knowledge bases [1, 58]. The importance of information seeking as a motivation for participation in online communities suggests that overlapping communities have a yet unexplored role in how people seek and find useful content.

2.2.2   *Homophily.* A specific type of need that online communities meet is to foster connections with similar others. We use the concept of *homophily*, "the principle that a contact between similar people occurs at a higher rate than dissimilar people" [54], to describe those benefits people can only receive in online communities of others who share their identities, beliefs, interests, or culture. In offline settings, homophily helps explain why tastes in cuisine, music, and other cultural preferences are so often correlated with each other and with political orientation [19] and why similar people tend to congregate and what happens when they do [54]. Homophily on social networks may drive the emergence of online "echo chambers" as individuals seek online communities whose members share their beliefs [20, 29, 34, 36]. People have greater degrees of trust in homophilous groups and are more likely to share content posted by homophilous others [14, 50]. Cultural homophily can be particularly important as demonstrated by how different language editions of Wikipedia vary enormously in the style of their contributions in ways consistent with differences in national culture [59]. Homophily also exists in online fandom communities where users gather to share their interests and experiences with media and entertainment [24, 33].

2.2.3   Finding a large audience. Research on online communities producing public information goods has found evidence that the size of the audience motivates contributors [78]. Additionally, numerous studies have shown that users of social networking sites are highly conscious of the audience that their posts and messages may reach [51, 76]. As individuals on social media typically have little information about who sees their posts, they conceive of "imagined audiences" based on cues from the visible activity of others [5], and target imagined audiences using deliberate strategies—such as using multiple platforms to reach distinct audiences—in order to control who sees or does not see their posts [48, 51, 79].

3   EMPIRICAL SETTING

To study overlapping membership in online communities, we conduct interviews with members of online communities hosted on Reddit, a social media platform for sharing, discussing, and rating news, media, and other content in user-created subcommunities called 'subreddits'. Individual users can participate in any of Reddits' hundreds of thousands of subreddit communities by either posting 'submissions' that might include a link to a news article, a question for discussion, an image to share, or new text written by the user. Users can also post 'comments' within the comment threads of posted submissions. Users vote submissions and comments up or down as a form of distributed moderation [42] and give them virtual awards.

Subreddit communities are managed by teams of volunteer content moderators tasked with structuring participation, curtailing abusive behavior, and keeping

other users on topic [53, 63]. Subreddits exist covering an enormous range of topics [23] and Reddit has been the site of much research on overlapping online communities [18, 31, 67–69]. The costs of creating and joining new communities on Reddit are very low, subreddits often overlap in both topic and membership, and users frequently create spinoff subreddit communities.

4 METHODS

To understand why so many people participate in multiple overlapping communities, we set out to interview people who are active in multiple highly related subreddits. Additional inclusion criteria were that users were adults (i.e., above the age of majority in their country) and able to participate in an interview in English. Our participation select process began by first choosing clusters of highly related groups. To do so, we referred to web-based data visualizations of clustering algorithm derived from topic and user overlaps to find groups of interest-based subreddits having similar users and similar language.

To generate the visualizations, we ran a computational analysis of the Pushshift Reddit dump [3], which contains a nearly complete history of Reddit comments as of April 2020. We selected the top 10,000 subreddits by the number of comments in this data, and excluded subreddits where a majority of submissions were flagged as 'NSFW' to avoid inspecting pornographic material. Next, following approaches in prior work on overlapping subreddits [18], we constructed measures of linguistic similarity and user similarity by taking the cosine similarities of TF-IDF vectors. To generate our two visualizations, we ran affinity propagation clustering [26] to

group subreddits having overlapping users or topics. We then built an HTML visualization of these clusters based on t-distributed stochastic neighbor embedding (t-SNE). We include these visualizations in the online supplement.

Although some aspects of cluster selection were necessarily arbitrary, we tried to select clusters that were interest driven, involved primarily English language discussion, and were focused on content and people that all members of the research team would be comfortable speaking with. As a result, we did not select any clusters that were focused on pornography or other content that is not safe for work, fringe or extreme politics, content specific to a geography, or topics that our group could not understand. We did seek out clusters that we hoped would result in individuals from a diverse range of ages, genders, and life experiences. The clusters we selected each include 3-10 subreddits on the following topics: rock climbing, streetwear fashion, roller coasters, vintage audio, podcasting, painting, drag culture and performance, indie music, and dating for middle-aged adults. Information about each subreddit and cluster can be found in Table 1.

Again using the Pushshift Reddit dataset, we selected candidate participants who were the top 80% of most frequent commenters within the cluster; who participated in multiple subreddits in the cluster; and were active in the cluster during a period of at least 1 calendar year. We began recruiting a random sample of 50 candidates matching these criteria within each cluster to participate in our study by sending direct messages through Reddit. Interested potential recruits filled out a short online survey confirming that they were adults and able to participate in English language interviews. We also asked participants about their participation and familiarity with each of the subreddits in each cluster to verify that they

Subreddit	Cluster	Subscribers	Created
r/bouldering	Climbing	194,814	2009-10-28
r/climbharder	Climbing	117,288	2010-10-19
r/climbing	Climbing	935,621	2008-07-17
r/climbingcirclejerk	Climbing	45,032	2011-08-18
r/Drag	Drag	44,724	2011-01-15
r/Dragula	Drag	27,510	2016-11-03
r/rupaulsdragrace	Drag	440,329	2011-11-15
r/RPDR_UK	Drag	31,867	2019-02-07
r/SpoiledDragRace	Drag	69,027	2018-02-16
r/MsPaintsArtRace	Drag	61,292	2017-04-17
r/MGMT	Indie Music	17,744	2010-02-25
r/tameimpala	Indie Music	94,248	2011-10-30
r/kgatlw	Indie Music	59,191	2015-07-01
r/Indieheads	Indie Music	1,932,698	2013-12-24
r/datingoverthirty	Middle Age Dating	436,480	2014-11-04
r/DatingAfterThirty	Middle Age Dating	11,550	2018-03-09
r/datingoverforty	Middle Age Dating	52,522	2018-12-15
r/relationshipsover35	Middle Age Dating	14,916	2018-02-06
r/OilPainting	Painting	186,716	2011-09-22
r/Painting	Painting	280,865	2008-06-13
r/PourPainting	Painting	178,800	2017-07-28
r/Watercolor	Painting	269,882	2012-01-15
r/HappyTrees	Painting	53,362	2011-02-07
r/podcasts	Podcasting	1,995,693	2008-01-25
r/podcast	Podcasting	60,497	2009-01-02
r/podcasting	Podcasting	73,010	2010-09-17
r/audiodrama	Podcasting	129,102	2010-11-30
r/ska	Podcasting	34,397	2008-03-12
r/guessthecoaster	Rollercoasters	5,094	2017-06-30
r/rollercoasterjerk	Rollercoasters	12,378	2016-07-14
r/rollercoasters	Rollercoasters	66,652	2010-07-31
r/rct	Rollercoasters	55,275	2010-08-04
r/themeparkitect	Rollercoasters	13,536	2014-06-16
r/streetwear	Streetwear	2,678,745	2011-04-30
r/supremeclothing	Streetwear	154,797	2012-04-04
r/womensstreetwear	Streetwear	421,279	2016-04-25
r/bapeheads	Streetwear	19,672	2013-08-12
r/malefashion	Streetwear	207,843	2011-04-02
r/sadboys	Streetwear	74,932	2013-06-30
r/techwearclothing	Streetwear	94,675	2017-03-01
r/Vans	Streetwear	51,997	2011-07-01
r/cassetteculture	Vintage Audio	45,615	2011-05-25
r/typewriters	Vintage Audio	20,037	2010-10-25
r/vintageaudio	Vintage Audio	59,202	2011-09-18

Table 1. Clusters of subreddits from which we recruited participants, subscriber counts at the time of the study, and the creation date of each subreddit.

were knowledgable. Although we recruited participants from the subreddits listed in Table 1, at the beginning of each interview, we asked if there were any other subreddits related to those identified by the clustering algorithm. As a result, our conversations were not limited to these subreddits.

We initially collected interviews from the first three clusters listed in Table 1. We found ourselves reaching saturation within these clusters more quickly than we expected. That said, we also found that different clusters were surfacing quite different data. In response, we sought to obtain better coverage of interest groups on Reddit by adding additional clusters and recruiting at least 2 participants from each. In some clusters added to the study later, we did not reach saturation in 2 interviews. In these cases, we sent additional batches of invitations and conducted additional interviews. In total, 20 participants were successfully recruited and interviewed by 5 different members of the research team before we reached saturation in our analysis and ceased the process of data collection. We describe the characteristics of our interviews in Table 2.

Our interviews were all semistructured. We drew from a long series of open-ended questions about participation in different subreddits and the relationships between communities but often chose our questions based on what our subjects wanted to talk about. A copy of our interview protocol is included in our supplementary material. Interviews were 49 minutes long on average but varied substantially. We suggested conducting interviews over Zoom, but also offered the participants their choice of communication channel. As a result, 2 interviews were conducted over the phone, 1 was conducted using Discord chat, and the rest were conducted over Zoom. Interviews were initially transcribed using Zoom's built-in

transcription and the otter.ai automatic transcription service and then manually corrected by the authors. After each interview, participants were compensated with a digital gift card for $20 USD through the Tango Card reward service[1].

5   ANALYSIS

Our analysis followed Charmaz's [15] approach to grounded theory as closely as possible. We began initial coding during our data collection period and generated over 950 codes which we then grouped in an iterative axial coding process that generated 18 distinct thematic memos. As we completed collecting data, we refined our codes and coalesced our themes to find an answer to our orienting research questions of "Why are there so many online communities organized around interest groups?" and "Why not more?" to identify three categories of benefits that interviewees seek from their participation in online communities and the sense that they look to different communities for different benefits. Additionally, our analysis was influenced by sensitizing concepts from prior work including the knowledge of overlapping online communities reviewed in §2.1 and the reasons that people participate in online communities summarized in §2.2. In synthesizing our grounded theory, we observed that interviewees described their participation in multiple different subreddits or their preference for particular subreddits in terms of the inability of one community (often the "main" or "largest" community) to provide the desired benefits.

---

[1]https://www.tangocard.com/

Participant ID	Cluster	Interview Length (minutes)
C1	Climbing	56
C2	Climbing	51
C3	Climbing	41
D1	Drag	51
D2	Drag	67
I1	Indie Music	71
I2	Indie Music	43
O1	Podcasting	30
O2	Podcasting	44
P1	Painting	58
P2	Painting	35
P3	Painting	40
P4	Painting	35
R1	Rollercoasters	24
R2	Rollercoasters	43
S1	Streetwear	79
S2	Streetwear	55
T1	Dating in Middle Age	63
T2	Dating in Middle Age	53
V1	Vintage Audio	34
V2	Vintage Audio	56

Table 2. List of anonymized participant IDs, the cluster from which we recruited them, and the length of their interview.

5.1 Ethical Concerns

Our study design was reviewed by the Institutional Review Board (IRB) at [RE-MOVED FOR REVIEW] and was determined to be exempt. As part of the design of this study, we took several steps to protect the privacy of our research partici-pants and their data. Participants were fully briefed about the design of the study before being interviewed and were given documents concerning the study as well as contact information for our IRB. Explicit consent was obtained from every sub-ject before recording interviews.

Because this project involved collaboration with a relative large team involved in data collection and analysis, we stored and shared interview recordings and transcripts using the Keybase end-to-end encryption service for all discussion and data sharing. Finally, participants were anonymized so that no direct identifier was

recorded in the process of data collection, and only anonymized pseudonyms (e.g,

C1, P2, V2 as show in Table 2) are published in this paper. *We made several minor edits to quotes to obscure potentially identifying details.*

6 FINDINGS

Why do people participate in multiple online communities around the same topic?
~~The~~ Our answer, ~~simply put,~~ *that emerged from our grounded theory* is that no one community can provide all the benefits *are that users want* that multiple online communities can. At a high level, we find that people have multiple and diverse motivations for participation in online communities. In §6.1, we describe the types of benefits they seek organized into three categories: (a) ~~to~~ engage with specific types of content; (b) ~~to engage in~~ homophilous socialization, and (c) ~~to share~~ content contributions with as large ~~of~~ an audience as possible. Using data from our interviews, we describe the tensions ~~between~~ between these benefits to explain why very similar groups of people talk about similar things in different online ~~communities~~ *spaces*. ~~The fact that online communities provide these three benefits may not surprise scholars of online communities.~~ Our contribution comes in the form of a theoretical framework, grounded in our data, that describes how ~~these~~ *the full* benefits *of participating in communities* can only be satisfied by groups of communities.

In brief, we find that a large community is valuable for its steady stream of content and a large audience, but participants find it more difficult to get attention and connect with others. Large communities tend to share content that appeals to a broader audience instead of the specific types of content that enthusiasts relish. Similarly, large subreddits provide less sense of community and fewer opportunities to meaningfully connect with others who share their specific hobbies and

interests. Although similar people often seek similar types of content, we find important cases where the desired content is often available only from dissimilar others. For example, beginners in sports like rock climbing or in art forms like painting often seek discussion with advanced practitioners to help them learn. No single subreddit can provide all these benefits. Instead, different subreddits provide distinct opportunities to find specific kinds of content, homophily, or the largest possible audience.

6.1 Benefits users seek from communities

6.1.1 Specific kinds of content. Content on Reddit is organized into subreddits that define boundaries around what is considered "on-topic." Multiple subreddits can exist around the same general topic when subreddits occupy different "conceptual rungs" and subdivide a general topic into specific subcategories like types of painting media, athletic subdisciplines, or audio formats. Moreover, subreddits that prohibit types of content or behavior open niches for subreddits that allow it. Despite such forms of specialization, multiple communities often welcome the same content and encourage users to "cross-post" material to support different and complementary discussions around the same topic.

Members reward content that fits within the category with upvotes and positive engagement. A subreddit's topic—what it is about, and what content should be posted—is often signified by its name. A climbing enthusiastic explains:

> I think the name itself [r/climbharder], kind of specifically points out
> that: This is not for people who climb hard. It's for people who climb
> and want to climb hard*er*. (C1)

C1 describes how the purpose of a subreddit is tied to its name by emphasizing the adjectival suffix '-er' as indicative of the fact that the subreddit isn't about achieving elite performance but about improving ~~even for beginners~~.

Similarly, a participant in subreddits about drag performance ~~and fandom~~ invokes Marshall Mchluhan to describe how they know what ~~messages~~ (content) to post and where to post them ~~through their knowledge of the media (subreddits)~~:

> ~~Yeah, well~~ let's say you were a drag artist and you wanted to show off something that you just created. You would have to go select which community you wanted to show it off in. And I guess among those, [r/Drag] would be the one to do that in. But if you're—if you're wanting to show off a piece of artwork or something that you made of an queen from Rupaul's drag race—and the best place to show that off would be to go to [r/rupaulsdragrace] and post it there. So it's more like [a] 'the medium is the message' kind of thing. … You know where where would get the most views [and] where would be the best place to post your content. (D1)

Like D1, our informants ~~All~~ had deep knowledge of what kinds of specific content would be appropriate for each subreddit ~~within~~ in their cluster. ~~We heard about subreddits that specialized in intersecting a topic with a dimension like nationality.~~ Specialization also occurred as a form of regulatory arbitrage when one community had formal or informal rules about the kind of content that was allowed. In these cases, we would often hear about an adjacent community where breaking the rules is accepted, ~~maybe~~ perhaps even the raison d'être. For example, r/rupaulsdragrace

prohibits spoilers and information about the outcomes of a reality TV show; r/spoileddrag is a community about the same show that allows spoilers.

This pattern is so widespread on Reddit that it is signaled in various ways. The "meta" prefix signals meta-discussions, often drama-centered, about another sub-reddit. The "jerk" suffix singals a space for memes, mockery, silliness, or other content unaccepted in the "main" subreddit. Both are commonly understood and were discussed at length by our interviewees. For example, among the Rollercoasters subreddits, the "jerk" subreddit is a "joke subreddit" where members of the main rollercoasters subreddit make fun of themselves:

> I would definitely say r/rollercoasters and r/rollercoasterjerk
> are really deeply intertwined. It's usually all the same members and stuff
> because of the fact that the coaster 'jerk' is just meant to make fun of
> the the main subreddit. It's just a joke subreddit. (R1)

'Jerk' subreddits were a common source of discussion among our participants. Similarly, among the Climbing subreddits, the "main" subreddit about rock climbing (r/climbing) is welcoming to newcomers. "People want to encourage more entrance into the sport" (C1) and so upvote posts by newcomers. However, newcomer posts are often repetitive pictures of people climbing in gyms or videos of famous climbers, and this annoys some experienced climbers. The "jerk" subreddit provides a backstage space where making fun of newcomers is acceptable.

In addition to being divided by formal or informal rules, interrelated subreddits can be structured as a ladder of "conceptual rungs" where one finds larger communities as one ascends the ladder. A participant in the subreddits on art and painting described this phenomenon as:

You kind of go up through these kind of conceptual rungs. … When you go up from, say, r/OilPainting—like r/HappyTrees to r/OilPainting— it's a much bigger community. And then from r/OilPainting to Painting, which is even bigger. (P2)

P2 explained that smaller subreddits like /r/HappyTrees support learners and are generally more welcoming places where the less skilled can find advice. This sentiment was also reported by Hwang and Foote [35] who describe small subreddits as existing within what they call "nested niches." While the quotation above suggests that the size of communities increases as one moves up the conceptual rungs, the relationship between topical scope and size was more complicated in our data than either this quote or Hwang and Foote suggest. In some topical areas, one finds that subreddits with relatively specific topics can have the largest or most active communities. For example, /r/rupaulsdragrace is the by a large margin the most active drag subreddit, even though it focuses on a reality TV series that is a subset of the broader drag community.

Although many subreddits with different specialized topics exist, people who want to share their work, ask a question, or have a specific discussion may not know the best place to post. Cross-posting—when someone posts the same content, questions, or messages in multiple communities—is a widespread act on Reddit that helps people find an engaged audience or discussion group. Our interviewees spoke about cross-posting at length. They described their motivation for cross-posting as the desire to increase attention to one's content, earn more reputation on Reddit (i.e., karma), obtaining more helpful or diverse information, and participating in different discussions about the same topic. Cross-posting has

sometimes been viewed negatively as a form of attention grabbing (i.e., "karma whoring"). More often, however, we heard that cross-posts ~~are~~ done to establish different conversations:

> They're going to get different responses from different subreddits. I think maybe some of them are kind of just posting to get extra 'karma,' but I think most of them, mostly it's just that you'll have a different conversation in the different subreddit. (V1)

Indeed, cross-posting to find different discussions or audiences in different subreddits is generally acceptable ~~as described by V1. It is frequently encouraged.~~ ~~Multiple interviewees~~ described how, when people ask for training advice in r/climbing, ~~the largest subreddit about rock climbing,~~ they will be advised to cross-post to r/climbharder~~, a subreddit specifically focused on~~ training:

> Somebody will post asking for advice in r/climbing and oftentimes, somebody will comment and be like, 'Hey, you know? You're ~~you're~~ welcome to ask this here. But you might get more and better responses at r/climbharder.' (C1):

C1 explained that even though conversations about training often start in the main subreddit, they are not likely to gain traction because not everybody in the main community is interested in the more intensive aspects of climbing. ~~When people start such discussions, they are commonly advised to ask again in a different community with a special focus on training.~~

In sum, the ecosystem of subreddits about similar topics provides more opportunities for people to find desired discussions. People receive positive feedback and

engagement when they post content that fits a subreddit's topic. Formal and in-formal rules prohibiting certain content or discussion allow for communities that embrace such content to grow in adjacent spaces. That said, the subreddit where a particular piece of content will be best received is often not clear to the person posting it. Cross-posting provides multiple chances to start a desired discussion.

6.1.2   Homophily. Online communities have long been recognized as a way to "find my people" by bringing together users sharing a psychiatric diagnosis, en-thusiasm for a hobby, age group, or membership in a subculture or identity group. A member of the Middle Age Dating cluster of subreddits explains:

> [When I joined the ADHD Reddit sites], I feel like I found my people
> after all these years. … If you don't have ADHD, and don't wonder
> what's going on other people's brains all the time, I think you just think
> that everybody thinks like you. And they don't. They don't. So if you're
> 30 and you're having a problem, you really just want to talk to other
> other 30 somethings. (T2)

T2's description of having "found my people" and talking to other people like themselves invokes the idea of homophily: the desire to connect to others similar to oneself. Analytically distinct from finding personalized information in narrowly focused subreddits, homophily was frequently cited as an end in itself by our in-terviewees. Our interviewees sought to connect with "like-minded" people having similar interests, demographics, identities, tastes, and levels of status.

Even though the identities of others in subreddit communities are largely invisi-ble, participants can easily imagine the demography of the subreddit. A participant

in the Drag cluster of subreddits described `r/Dragula`, a community of fans of a TV show featuring horror-infused drag styles, as:

> I think it would be a mostly LGBTQ audience. And not many straights. But if there are straights, they would be really open minded or edgy. Or, I don't know … associated with that 'dark' aesthetic. (D2)

D2's thoughts on `r/Dragula` convey a clear sense of the audience the subreddit. Of course, pseudonymous nature of Reddit obscures age, race, gender and ethnicity. That said, Reddit users draw on stereotypes about fanbases and cues like mentions of schools, selfie posts, linguistic markers, and cultural references, to build clear models of the types of people in a subreddit. In further unpacking these dimensions, D2 contrasts `r/Dragula` with the more mainstream subreddits about the show "Rupaul's Drag Race:"

> [As for subreddits about] the drag race (`r/rupaulsdragrace`), Drag Race UK (`r/RPDR_UK`), and the spoiled drag race (`r/SpoiledDragRace`). … Most of [followers in these other groups] don't do drag. Most of them are, I think, white gay men, or straight women who see drag with a very narrow view of what drag is. Hegemonic? I don't know if that's the word, but they apply the same standards of beauty that are applied to women and men and artists and performers to this art form. And it limits a lot what you would consider pretty, or you would like, or you will enjoy. (D2)

D2 conveyed both a strong sense of the demographics of different drag subreddits as well as a strong sense of identification with `r/Dragula` which they described as less toxic, more inclusive, and more creative, in part because its membership has

a greater concentration of LGBTQ and nonwhite people who are less interested in conforming to hegemonic beauty standards. The mainstream drag subreddits are prone to reproducing dominant beauty standards, but `r/Dragula` provides an alternative space for drag queens and kings who do not conform.

Subreddits divide broad topical areas like drag, art, and fashion into subgroups of people occupying strata of status hierarchies associated with identity, expertise, and class. In the Climbing cluster of subreddits, rock climbing ability confers status and separates beginners from advanced athletes. We found that these two groups concentrate their participation in different subreddits. Across the clusters, we found that experts sought out fellow experts with whom to share knowledge, offer reflections, and give advice grounded in shared extensive experience. We found that status hierarchies in Painting subreddits are related to skill level and medium:

> I do believe the moderator had asked me if I would post [on] `r/Artoilpainting`. But it seems like they frown on watercolor. So I kind of I just kind of don't post there. ... Because oil painters like to be snooty, because oil painting is expensive. ... They like to look down on watercolor. (P4)

P4's story of status-based divisions between groups was a consistent theme across clusters in our analysis. In this case, P4 describes how they were invited to cross-post their work from `r/Watercolor` to `r/Artoilpainting`, a smaller subreddit that seems to have a complicated relationship with watercolor. While watercolor submissions are allowed, and in this instance encouraged, both the subreddit's name and the semblance of its visual tag for watercolor submissions with the downvote button, suggest that oil is the preferred medium in this community.

Similarly, our subjects reported subreddits about fashion are split along lines that are associated with the price and status of the clothes being discussed:

> The kind of person, the platonic ideal poster, like user, of something like `r/streetwear`, is probably more open-minded, maybe, in terms of what they think is cool, what they think is worth wearing. Whereas, you know, . . . the `r/malefashion` snob is a snob. (S1)

Even though users of `r/streetware` share and discuss men's fashion, their casual and youthful styles are looked down upon in `r/malefashion` which focuses on higher-status and more expensive styles. S1 is a member of the `r/streetwear` subreddit. They identify the `r/streetwear` poster in positive terms—e.g., as having an "open-minded" approach to fashion compared to the snobby `r/malefashion` members. While their groups are "chill" and "supportive," higher-status groups are "snobby." It is clear that S1 feels unwelcome and out of place in the higher-status group. In this way, the division of topical spaces into spheres of similar status and identity allows members to find groups that exclude those who look down upon them as well as those who they look down upon.

Although "finding your people" is satisfying in itself, it can also be a foundation for a wide range of other kinds of benefits. For example, a community leads to conversations that are more personal and promotes trust. Trust, in turn, is useful for buying, selling, and trading and builds confidence in the advice and information shared within a community.

V2, one of our interviewees from the Vintage Audio cluster, described a community of record collectors on Reddit that acted as a market for buying, selling, and trading records. They preferred this subreddit to other online markets like

Ebay because the community holds members accountable for honest transacting and because of the intrinsic reward that comes from sharing records with a fellow community member:

> because it's a group of people that are like-minded, . . . your feet are kind of held to the fire a little bit more about actually being realistic with the condition [of the material you are selling]. Whereas, [when you buy] vinyl at the used record shop, sometimes you feel like someone's trying to pull one over on you . . .
>
> I feel like because it is a community, sometimes you can get some kind of better deal. Because it'll be like, hey, if somebody really liked, whatever, Modest Mouse, I've got these records that I want to sell to somebody that likes Modest Mouse instead of putting them up on Ebay. . . .
>
> I found other people that share the hobby that I like. So I almost, definitely, feel like they're friends in a little way. And so I want to, if I'm ever selling, I'm going out of my way to make sure that whatever I'm doing, everything I'm doing, is above board. (V2)

V2 was very enthusiastic about the "marketplace wrapped in a community" for vinyl records. According to V2, both buyers and sellers of records benefit from transacting within a community of like-minded hobbyists. Because the community holds sellers accountable through public shaming ("holding their feet to the fire") and the threat of exclusion from the market, the community promotes honest representation of merchandise. Being part of a like-mindeded community where members feel friendship with each other gives sellers a reason to be honest, and

even to discount their wares, because they get "some kind of better deal" on Reddit than on another market.

In sum, our interviewees turned to specific subreddits to find people who share their interests, tastes, problems, and identities. Our participants described subreddits in terms of demographics and identity groups as well as well as styles, subgenres, or categories related to social status like wealth, expertise, and beauty standards. Universally, they used these categories to place themselves within the constellation of related subreddits they participated in. Members of subreddits who are "finding their people" benefit one another by acting as communities as well as building trust and feelings of friendship. Over time, these feelings can provide further benefits like the ability to more safely engage in buying and selling.

6.1.3 Finding the largest possible audience. A third type of benefit derives from the number of members in a subreddit. Our interviewees were all keenly aware of the fact that a post reaching one of the top positions on a larger subreddit would receive the attention of a vast audience. They described this attention as emotionally thrilling and as translating into other benefits. For artists and influencers, large audiences brought material rewards. For learners, a large audience's collective knowledge could bring hard-to-find answers and advice.

That said, our interviewees explained that larger subreddits do not necessarily provide a larger audience because posts in larger subreddits are more likely to be ignored or missed in the torrent of other content. While posting in a smaller subreddit might increase the chances of finding an audience at all, subreddits that were too small were described as unattractive because they would not attract many

posts or replies. Interviewees responded by choosing where to post strategically to find the largest possible audience.

Although the competition for the top spots on the front page of large subreddits can be fierce, this competition can make recognition from a large subreddit extremely gratifying:

> Likes are just kind of fake: fake social currency. But yeah, when you get a charge out of it, yeah, I love it. Most of the time, painting is a really busy sub. I mean, like, in any given hour, the new page is already replaced. …
>
> If you can get something that gets a hold there and stays on the front page for a little while, [if] it gets up in even the top five, I've had a handful do that. That's kind of cool. (P2)

P2 describes the thrill of reaching top positions in r/painting with posts of their paintings. Even though they are dismissive of likes on Reddit, they admit that the attention their work gets from the subreddit is desirable. It sends traffic to their websites, raises their artistic profile, and helps them sell their art. While these material incentives are important, part of the thrill comes from knowing that a given subreddit is competitive. Smaller subreddits are simply unable to provide these benefits.

While small subreddits are ineffective at providing attention or information, posting in a large subreddit means the risk of being ignored:

> It's hard. I think there's this weird bell curve where the community needs to be big enough where people want to post content. But it can't get too big where people are drowning each other out for attention. (S2)

S2 was among several of our interviewees who described an ideal "middle ground" for subreddits size. In general, we heard that people were less interested in posting content in very small subreddits that do not provide an audience. That said, competition over the largest audiences drives people to smaller subreddits where they can reliably find an audience. I2 from our Indie Music cluster explained:

> Usually r/Indieheads is the way to reach more people if you want to. Just like if you wanted to do even more, you'd probably do it on r/music. … Say a small indie band decided to do an AMA they would probably want to do it on r/Indieheads. Because if they did it on r/music, it would get drowned out and nobody would see it because there's so many posts. In r/Indieheads it would get a decent bit of attention, I think. In the band subreddit, it would probably get a lot of attention too. But r/Indieheads seems like the best middle ground for that kind of thing. (I2)

I2 explained that when the psych-rock band *King Gizzard and the Wizard Lizard* wanted to engage with an audience on Reddit, they had a choice whether to post in either the smaller "band subreddit" dedicated to them, the very large r/music, or the medium-sized r/Indieheads. Although posting in the band subreddit would have surely provided an audience, they chose r/Indieheads which was large enough but where there was still little risk that their post would be drowned out.

Our interviewees repeatedly described how finding an audience for one's content is a clear motivation for posting in larger subreddits. However, we also heard that competition for attention in the largest subreddits leads people to try to find an audience in smaller subreddits. As one posts in increasingly small subreddits,

one finds increasingly small audiences. In the smallest subreddits, posting may not seem worthwhile at all. This trade-off between finding a large audience and being ignored suggests that posting in subreddits of intermediate size can be the most reliable way to reach a sizable audience.

6.2   Tensions between the benefits

The sections above describe the three broad classes of benefits that users sought out and how this affected the types of subreddits that people joined to find these benefits. How does understanding these benefits help explain why it is so common for individuals to belong to multiple groups about similar topics?

The findings in the previous sections imply a clear reason that so many overlapping subreddits exist. When one subreddit prohibits a certain type of content or conversation, an adjacent group can form that allows it. When an identity group is marginalized in one subreddit, members of that group may form a subreddit of their own. When getting attention in a large subreddit is too difficult, a smaller subreddit becomes attractive.

However, this explanation does not directly engage with the evidence in our interviews of real tensions between the three types of benefits. Using data from our interviewees, we describe each of the three possible tensions that exist between the three benefits: (1) subreddits where one finds a large audience are less able to provide specific types of content; (2) communities with large audiences are rarely able to provide a community of similar others. While people who share an identity sometimes seek the same types of content, (3) some valuable types of discussion and information only are found in diverse groups of people. As we

discuss in §??, taken together these tensions have the a form of "trilemma"—i.e., a dilemma between three mutually incompatable options—existed between our interviewees' desires for specific content, homophily, and finding audiences. A single community might provide 2 of these benefits. But almost never all three.

6.2.1 Larger audiences create background noise. In §6.1.1, we described how subreddits are structured according to distinctions between different types of content. Breaking down broad topical areas into specific subreddits of varying levels of granularity makes finding such content easier because it means that users do not need to sift through unrelated material in a broader subreddit. One common pattern in our interviewees was the idea that larger subreddits are simply not the best place for enthusiasts to have discussions:

> I see this background noise problem building [in] r/climbing, the main climbing community, [which] has just become less and less and less interesting and less relevant as it's gotten bigger. That's not really a problem. Right? That's probably has more to do with my interest level and how long I've been on it. And my experience level with climbing. I'm just a little bit more crusty about it, you know? (C2)

C2 describes losing interest in the primary subreddit about climbing as it grew because of the interviewee's specific interest in particular types of climbing content (i.e., material associated with being "crusty" or experienced). C2 recognizes that when r/climbing experienced growth, it didn't lose its character as a community. That said, a larger volume of posts by newcomers to the sport created a "background noise problem" that made it difficult for established climbers to find

discussions of interest. C2 recognizes that this is due to their own increasingly idiosyncratic interests.

Similarly, smaller subreddits can be incredibly valuable to those looking for highly specialized information. Even though they may have very low levels of activity, they can provide a way to learn about rare forms of expertise. A participant in our Vintage Audio cluster explained how they might seek out advice on building an reel-to-reel audio setup:

> If you're at [r/ReelToReel]. Everybody is hyper into them. Whereas there's probably overlap with somebody in r/vintageaudio ... For me, if I'm like trying to rebuild my reel-to-reel player, I want to talk to that kid, or 2 or 3 kids, 4 kids, the most knowledgable person particularly about building reel-to-reel.
>
> And if there wasn't that reel-to-reel subreddit or a "building reel-to-reel" subreddit. You know what I mean? Like, I don't feel tempted to try that out. So I know that who I'm talking to is hyper specific to the knowledge I want. (V2)

Invoking r/ReelToReel, V2 describes a highly niche subreddit about archaic audio tape equipment with only 3,200 subscribers and a handful of posts each day. Regarding specific technical information, V2 is simply not looking to find a large audience. Instead, they want access to the "most knowledgable person" with specific expertise because access to this expertise makes it possible for them to consider doing their own reel-to-reel projects.

Although the `r/ReelToReel` community overlaps with the larger and more general `r/vintageaudio`, the latter does not provide the ability to connect with a small group of expert enthusiasts in an archaic and obscure technology.

Similarly, when someone wants a podcast recommendation tailored to their personal tastes, asking in a larger subreddit is not likely to prove as fruitful as it is within a smaller one. O2, a participant in the Podcasting cluster explained:

> So I think for like `r/audiodrama`, I would probably write a longer post, and probably get a bit more into like, my personal tastes. Like I would comment about, 'oh, I really love the acting in this one, is there anything similar?' Open up a bit more about what I do and don't like. Whereas I think in podcasts, it probably would be more direct. I'd ask a specific question … more to the point, more factual, probably just more almost transactional. (O2)

Although the larger `r/podcasts` subreddit is a popular place to promote podcasts on Reddit, O2 explains that they prefer asking for recommendations in the smaller `r/audiodrama` where they find others willing to take their personal tastes into account. Our interviewees did not advance a "smaller is better" argument. O2 explains that they still engage in larger subreddits, but that they approach it as a more direct and transactional mode of information exchange when they do. Similarly, large art communities provide opportunities to find a large audience, but when someone wants substantive feedback to improve their skills, they post in a smaller subreddit organized specifically for this purpose.

Interviewees described the most general interest-based subreddits like `r/podcasts`, `r/painting` and `r/climbing` as more accessible and welcoming to newcomers

and as reaching a larger audience. All things they valued in certain circumstances. They also described these larger groups as having a high volume of low-effort posts or comments. Our interviewees explained that although they play a useful role in an information ecosystem, the largest subreddits in a topical area are rarely the best places to look for information or advice. They explained that small subreddits exist because they can effectively play host to content, information, and discussion that larger subreddits can not.

6.2.2   Homophily is more difficult in larger groups. Because they have less background noise, smaller subreddits are more likely to provide opportunities to connect with people who share one's distinctive interests, tastes, and identity. Smaller subreddits are also better places to find a community because they provide opportunities to have repeated encounters with recognizable others, off-topic discussions, and more personal interactions. P4 explained that building relationships was difficult with large and broad audiences:

> Obviously, I want as many people to see my stuff as possible, especially
> [since I am] trying to establish myself. But at the same time, I do want
> to build a relationship with any sort of community that I can. (P4)

P4 explained that they participated in multiple communities because they have two goals as an artist. First, they want to find an audience for their artwork to establish their career. Second, they want to build a community with others who share their craft. They felt that they needed to turn to multiple subreddits to fit these different needs.

Although larger subreddits provide a large potential audience, smaller subreddits were described as being friendlier. Another interviewee from our Painting cluster explains that this is because of how people act differently in large and small subreddits:

> I live in the middle of nowhere. And every so often, before the pandemic, I would visit the [large city several hours away]. Now I found there were very polite people, both in [the city] and in [my rural area]. But the tone by which people carried themselves changes in their environment: that's kind of one of the big changing factors. So, in the city, people are in a rush, they're about their business. We don't really have time to chat. … The big subreddits might seem unfriendly [But] it's not that so much. Individual members are impolite or unfriendly. But it's almost as though people carry themselves differently when we're in different subreddits. (P3)

In their extended metaphor, P3 explained that large subreddits are like big cities full of busy people who do not "have time to chat." Evocatively, they described people as carrying themselves differently in large and small subreddits as their behavior changes when they move between different environments. The very same people who are rude in large subreddits might be friendly in smaller subreddits where people have repeated encounters with one another and have a stronger sense of knowing each other.

In another quote from the same cluster, P2 described how the small subreddit for Bob Ross inspired painters, `r/HappyTrees`, stands out from the larger art subreddits because people know one another and it does not feel anonymous. The

tight-knit nature of this community contributes to its utility as a source for feed-back.

6.2.3 Tension between finding specific content and homophily.
A third tension de-scribed by our subjects was between the desire to find discussion, content, and feedback and the desire to find similar others. Our interviewees described a range of situations when they sought out dissimilar others. For example, they described beginners seeking to learn from experts and outsiders seeking to learn about other cultures. They also described how subreddits instituted rules to limit or orga-nize content that also interfered with unstructured and off-topic discussions that helped with community building.

For example, although multiple subreddits with overlapping users discuss the same episodes of the TV series *Rupaul's Drag Race*, they have different understand-ings of events in the show depending on their national identities. D1 explained that:

> The discussions played out differently on different subreddits. In the Drag Race UK sub there's a lot more understanding about [a British drag queen] in particular, about where they come from … In America we don't understand how that person is from Worcestershire. (D1)

D1 explained that the cultural background of one of the drag queens was a sub-ject of discussion in r/RPDR_UK, the UK drag race subreddit while the main subred-dit, r/rupaulsdragrace, was "dominated by the American viewpoint" and lacked these discussions.

Our interviewees described a number of subreddits focused on discussing broad topics from a specific national or regional culture context. These cultural communities within a topical area provide a homophilous space for sharing distinctive cultural knowledge and sensibilities.

An additional wrinkle is that even for our American interviewee, the r/RPDR_UK subreddit provided an opportunity to enhance their own experience and appreciation of the show by observing and learning from members of another culture. In examples like these, our inteviewees explained that communities where like-minded people can share their distinctive appreciation show could provide not only opportunities for homophily by insiders but also a unique source of knowledge for outsiders.

Similarly, Painting participant P2 explained that a group that has a mixture of experts and beginners provides a better learning environment than a group of beginners alone:

> If you can find a small group, with a small core of people who are particularly skilled, they sort of energize the group as a whole. r/HappyTrees, even though it's kind of a beginner subreddit, there's some people that posts there that are like, you know, Bob Ross instructors, or they've been doing this for years. And they've mastered that sort of . . . "happy trees" thing. (P2)

P2 explains that part of what makes r/HappyTrees great is that it connects learners to a group of experts. A homogenous subreddit of only beginners or experts would not provide the same opportunities for new painters to get help with

their work from more experienced ones or for experts to give advice to beginners. For obvious reasons, this is not possible in a strictly homophilous group.

To stay focused on specific types of content, subreddit moderators will frequently employ stricter rules and heavy-handed moderation. A participant in the Indie Music cluster explained:

> A lot of them put in rules now where it's like, you can only do one post today, or one post a week, to try and limit that spam of people posting their own bands. ... In r/Indieheads, you can't post yourself playing music. ... They're kind of strict on the rules. Like you have to have your post titled a certain way to say whether it's a new ... single, sorry, a new music video, or new album. You have to like put it in brackets like [album], or like [new live stream]. And every post I think has to be like tagged like that. (I1)

I1 explained that the subreddit r/Indieheads has rules limiting how often one can post, requiring specific titles and tags, and prohibiting types of user-generated content. While these rules help maintain a high-quality feed of specific types of content, they can get in the way of building a community. As a result, subreddits that make rules to ensure that posts are high quality and on-topic frequently have adjacent "-jerk" subreddits that provide an outlet for jokes and memes and act as places where off-topic discussions and chit-chat can thrive.

Our respondants explained that smaller subreddits can get by with fewer rules and lighter moderation because they have fewer behavior problems and are less attractive to toxic outsiders. They are also more able to self-police using Reddit's

voting system and through direct interpersonal sanctions like admonition. One of the Vintage Audio participants explained that:

> In Reddit, the more users you get, the more strict the rules, and the more strict the moderation. Just to prevent problems. (V2)

V2 continued and explained that subreddits that are small enough that you can "wrap your hands around" and that are built around a "like-minded" group and have more tightly knit communities can develop and enforce shared behavioral norms that substitute for formal rules and rigid enforcement regimes. In concrete terms, V2 explained that the processes of creating spaces for specific types of information got in the way of building community. In smaller subreddits, when members want to start a conversation about something, even if it is silly or off-topic, they can do so.

7 DISCUSSION

How do overlapping communities come to exist ~~in the first place? Why do people belong to multiple groups about highly overlapping topics?~~ Our findings offer one answer. *When* Individuals seek multiple kinds of benefits ~~from online communities~~ that cannot be provided by a single online community, *a* An ecosystem of overlapping ~~online~~ communities ~~exists~~ *emerges* to provide the full range of benefits that users seek.

7.1 The Specificity-Homophily-Audience Trilemma

The tensions between the benefits that our interviewees sought can be thought of as forming a 'trilemma' between finding specific content, homophily, and finding an audience *as long as* *as possible*. This ~~description of a~~ three-way dilemma captures the fact that the
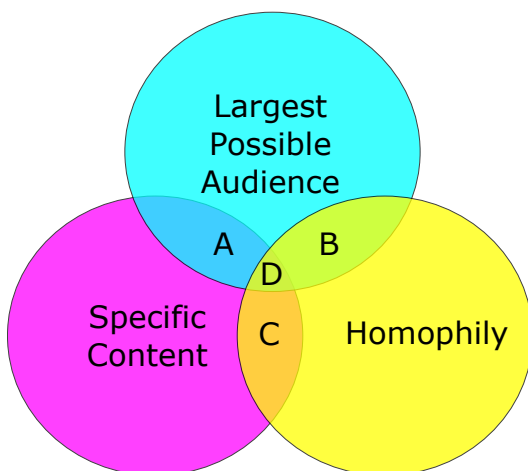
Fig. 1. Venn diagram illustrating the specificity-homophily-audience 'trilemma.'

more a subreddit succeeds in providing one of these benefits, the less able it will be able to provide the others. Subreddits might be able to provide two of the benefits to an individual. That said, only in truly exceptional cases will they be able to provide all three. Because communities have to structure themselves in one way or another, users opt to participate in multiple communities to gain all three benefits.

This trilemma is visualized in Figure 1. Each of the benefits described in §6.1 is reflected in the large circles. Each of the tensions described in §6.2 is reflected in the overlapping areas in the figure. Area A reflects how a community that overlaps in a large audience and specialized content may be unlikely to provide homophily to community members. Subreddits that provide large audiences face "the background noise problem" as a large volume of submissions makes it difficult for people to find the content they care about the most. Area B reflects the fact that a community that overlaps in a large audience and homophilous community will struggle to provide specific content like how an American interested in learning about international drag culture finds the need to search beyond r/rupaulsdragrace.

41

Area C attempts to represent how a community that overlaps in specialized content and a homophilous community will be unlikely to attract a large audience. Although not everyone who desires a specific type of content may be similar to those smaller who produce the content, smaller subreddits can often provide both desired content and opportunities to socialize with similar others. However, as the size of the audience increases, subreddits encounter the background noise problem and acquire a "big city" air of unfriendliness.

One might call this three-way conflict a "soft trilemma" because perhaps in exceptional cases an online community may be possible in Area D, where it is able to provide all three benefits to an individual community member at once. Additionally, our interview data suggests that the tension between two homophily and finding specific content (Area B) is weaker than the tensions involving audience size. Hard or soft, our findings suggest that individual subreddits are constrained by these trade-offs described by our interviewees.

7.2 Connections to Prior Research

7.2.1 Finding specific content. Our findings are consonant with prior work that ranks entertaining and informative content among the primary benefits that provided by online communities because stem from their power to connect people to novel and hard-to-find information sources [4, 9, 24, 72]. Within the interest-based subreddits our interviewees participated in, they often sought out conversations and discussions in the comments. Our study adds to contributes to this literature by showing how nested

and overlapping online communities are useful for information ~~search~~ *seeking* and the control of one's information exposure. Individuals often desire multiple types of content within a general subject area—~~sometimes explicitly at different times as~~ *conflicting* ~~with spoiled and spoiler-free discussions—that they can find in different overlapping communities~~ *like*. Even when an ~~online~~ community *relative obscure* like `r/vintageaudio` exists, ~~around a general topical area~~, an even more specialized community like `r/ReelToReel` may provide access to an even more specialized set of experts.

7.2.2 Homophily. Prior work has recognized the importance of homophily in motivating and structuring participation in online communities [14, 16, 29]. Contributing to this line of research, we identified a number of types of homophily that ~~drove~~ *drive* individuals' decisions to participate. These included hobbies, expertise, age, national culture, identity, and status. ~~However,~~ we ~~also found cases where~~ *homophily was in tension with the need for specific content in that* differences along many of these dimensions were valuable, ~~especially~~ for finding information.

Our results
This suggests that participants in online communities face trade-offs between homophily and information novelty. These ~~are~~ *maybe* similar in structure to ~~those~~ *the tradeoffs* between ~~close~~ *short* and long ties observed in contexts like work groups [62] and ~~online and offline~~ social networks [28, 29]. One advantage of joining a group of overlapping online communities is that it can help users to find information that would be unavailable in homophilous groups.

7.2.3 Finding the largest possible audience. Much social computing research points to the benefits of large audiences and large communities [39]. Our work adds more evidence to back up those claims. More relevant, perhaps, are recent counterclaims

about the benefits of smallness. In particular, a new paper by Hwang and Foote [35] presents an interview study with members of small Reddit communities. Al~~though~~ *(While)* our results about the tensions between ~~largeness~~ *audience size* and other benefits are fully in line with Hwang and Foote's findings, our starting assumptions and ~~our~~ ultimate takeaweays are quite different. Hwang and Foote seek to understand why people participate in persistently small communities and conclude that smallness offers a range of benefits. Our results suggest that individuals seek out benefits that happen to be incompatible with largeness and participate in portfolios of communities that, due to the trilemma we describe, will almost certainly include small ones. Although we believe that Hwang and Foote's [35] emphasis on smallness might draw focus to a side effect instead of the cause, we believe that the findings in our two papers are largely complementary.

Our findings provide ~~additional~~ evidence that although users may desire large audiences, large online communities often require additional structure to maintain order. Kiene et al. [38] describes how a massive influx of newcomers presents difficulties that can be managed by appointing additional moderators, increasing norm enforcement, and limiting the frequency of posts. Lin et al. [47] find that such interventions help subreddits maintain comment quality and stay on topic during massive influxes of growth. ~~However,~~ *Our sense is that* these changes ensur~~ing~~ *e* the availability of specific content ~~are~~ *is* effective, in part, because of the growth-limiting effects of rules and enforcement [30, 70]. *We see this as yet more evidence in favor of our theory.*

7.3   Implications for Ecological Studies in Social Computing

With the exception of a small qualitative subpart of a paper by Zhu et al. [80], prior ecological studies in social computing ~~on online community ecology~~ have relied on concepts like competition and mutualism but have provided limited evidence that such concepts are salient to ~~members of online communities~~ participants. As part of our interviews, we asked our interviewees if they perceived relationships between the communities they participated in to be competitive or mutualistic. In some cases, interviewees imagined hypothetical scenarios where competition might emerge from the perspective of subreddit moderators. For example, a participant in climbing said:

> I guess if you put your Reddit [moderator status] on your resume or
> something, and you want to be a moderator of a larger community, you
> could try ~~to try~~ to get users from other communities. But I haven't seen
> or experienced competition. (C1)

Although we asked nearly every interviewee about competition, only one interviewee (S2) described an actual instance of conflict or direct competition ~~that led to the creation of a a rival subreddit~~. In nearly every other interview, our subjects found our ~~implication~~ suggestion that subreddits ~~could~~ might be in competition to be suprising and ~~foreign~~ strange.

On the other hand, the idea that ~~they~~ communities are complementary ~~or~~ and mutualistic was much more intuitive. One Vintage Audio participant explained the relationship between subreddits:

> Yeah, the overlapping. … They each have their own niche. … They get
> big enough to have super critical mass of people. Then they'll have a

reason to exist. And then they'll sort of fit into the ecosystem of different

communities. (V2)

In ways that tracked this ~~rather~~ unproblematic description of subreddits coexist-

ing, our interviewees repeatedly suggested that the lack of structural or technical

limitations to the number of subreddits one can join reduces the possibility of

competition.

7.3.1 Resource Partioning.
The quote above by V2 can be ~~interpreted~~ as a kind

of summary of resource partitioning theory (RPT), another strand of ecological

research in organizational science that ~~considers the role of~~ specialization [10].

Although RPT has not been deeply examined in prior social computing work, our

findings suggest that it may be able to explain the widespread occurrence of over-

lapping communities. RPT proposes that the reason that small ~~and more~~ special-

ized organizations coexist with large generalist organizations is that generalists

are constrained in their ability to meet distinctive needs in niche markets [11, 66].

In V2's terms, the "ecosystem of different communities" is constructed by a pro-

cess in which those that "have a reason to exist" and specialize to "have their own

niche" and achieve "critical mass."

Our grounded theory suggests that the trade-offs in the capacity of an online

community to provide different types of benefits that people seek from online

communities give rise to new niches ~~within topical areas~~. Based on our findings

and our understanding of RPT, we hypothesize the following process to describe

how systems of overlapping communities develop:

When a new topical area grows, the vast majority of activity will happen in a "main" community. New members joining that community may seek and find the perceived benefits we described in §6.1 (specific kinds of content, homophily, and the largest possible audience) for their own content. If the topic area, like "art," is sufficiently general, then initial membership growth occurs as the community attracts new and existing Reddit users interested in both general and specific types of content and discussion.

As growth continues, membership in the generalist community becomes heterogenous in dimensions of homophily like expertise (e.g., amateur and professional artists) in specific interests (e.g., painters and photographers) and in the types of engagement desired (e.g., attention from an audience or critique). At this point, as we discuss in §6.2, the trade-offs related to size become relevant. In other words, finding information related to a specialized subtopic within the general topic and socializing with similar others each grow difficult in the growing community.

If, as with Reddit, creating new communities is virtually costless, a community specialized in a subtopic of the general topic can emerge. This specialized community will likely not attract as large an audience as the generalist community, but those most interested in the specific subtopic will join it to escape what our interviewees describe as "background noise" in the larger generalist. Similarly, those seeking personal interaction or social bonds with other community members will be more likely to find them in the specialized community. Subtopics are not the only way for communities to specialize. A similar process occurs in the formation of spaces having different rules or purposes (like "jerk" spaces). The cycle will then

begin anew as subreddits repartition a subtopic like `r/painting` into subspecialists like `r/oilpainting` and `r/watercolor`. Although some of our interviewees theorized such processes involving the emergence of new specialized communities from large ones, the model we have narrated is an untested theory. We leave it to future work to establish its empirical validity.

7.4 Implications for Design

The clearest implication for design from this study is that no single online community can do everything. By allowing users to create multiple communities having similar or identical topics, platforms can host ecosystems of online communities capable of providing a greater range of benefits to a wider range of users. Our findings also suggest that supporting ecosystems of overlapping communities is an important and perhaps overlooked design space.

In describing their engagement with multiple overlapping communities, our interviewees discussed affordances and limitations of Reddit's design. They found that it was sometimes difficult to find communities aligned with their interests. They explained that Reddit's system for recommendeding subreddits returned suggestions that were not relevant and its "multireddit" feature for making a custom feed of subreddits had an overwhelming number of options. While some of our interviewees reported discovering niche subreddits through these features, they also often found them through search engines, links in comment threads, and suggestions for cross-posting from other users. Many Reddit users make heavy use of the aggregated streaming feeds `r/all` and `r/popular` which surface highly upvoted posts from across Reddit. However, these feeds typically feature content from

subreddits that are already extremely popular instead of smaller subreddits with narrow topics matching enthusiasts' deep interests, expertise, and social groups. There remain many design opportunities for tailored suggestions of subreddits based on the mix of benefits described here.

7.5 Limitations

Our study has limitations common to all interview-based studies. Our findings derive from in-depth conversations with relatively few of the people who were highly active participants in the handful of clusters of communities from which we drew our sample. Although our study was designed to provide saturation within each cluster and to cover a wide range of types of clusters found on Reddit, additional interviews across a wider range of communities might uncover new types of specialization. Additionally, our interviewees were all among the 20% most active members of the clusters we sampled from. Additionally, our interview data were collected at one point in time and cannot speak to how the dynamics we describe played out over time and new communities were created and emerged. Similarly, although we find that overlapping communities tend to provide different benefits to members, we did not set out to interview community founders and thus cannot speak to the reasons that communities were created [25].

Furthermore, our study focuses on the Reddit platform. Although Reddit is among the most popular sites for online communities and we believe that the clusters from which we draw our interviews are typical of online communities organized around interest groups, our findings may not generalize to communities on other

platforms. Reddit has distinctive affordances provided by systems for voting, moderation, and multicommunity engagement that might shape the construction and use of overlapping communities. Our findings may not generalize to how overlapping communities on other platforms or between one platform and another may have relationships.

8 CONCLUSION

Why are the same people talking to each other in different online communities about similar things? We answer this question by developing a theory grounded in the analysis of 20 interviews with members of highly related communities on Reddit. Our answer suggests that people turn to online communities in search of multiple different benefits—specific kinds of content and discussion, socialization in a homophilous community, and attention from the largest possible audience. Although structures like the topic, rules, and size of a community might improve the degree to which it provides one of these benefits, they also detract from its ability to provide others. Within a topical area, multiple communities having a range of structures exist to provide the full range of benefits.

REFERENCES

[1] M. S. Ackerman and T. W. Malone. 1990. Answer Garden: A Tool for Growing Organizational Memory. In *Proceedings of the ACM SIGOIS and IEEE CS TC-OA Conference on Office Information Systems (COCS '90)*. ACM, New York, NY, 31–39.

[2] Lada A. Adamic, Jun Zhang, Eytan Bakshy, and Mark S. Ackerman. 2008. Knowledge Sharing and Yahoo Answers: Everyone Knows Something. In *Proceedings of the 17th International Conference on World Wide Web (WWW '08)*. Association for Computing Machinery, Beijing, China, 665–674.

[3] Jason Baumgartner, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn. 2020. The Pushshift Reddit Dataset. *Proceedings of the International AAAI Conference on Web and Social Media* 14 (May 2020), 830–839.

[4] Yochai Benkler. 2006. *The Wealth of Networks: How Social Production Transforms Markets and Freedom.* Yale University Press, New Haven, CT.

[5] Michael S. Bernstein, Eytan Bakshy, Moira Burke, and Brian Karrer. 2013. Quantifying the Invisible Audience in Social Networks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 21–30.

[6] Petter Bae Brandtzæg and Jan Heim. 2008. User Loyalty and Online Communities: Why Members of Online Communities Are Not Faithful. In *Proceedings of the 2nd International Conference on INtelligent TEchnologies for Interactive enterTAINment*. ICST, Cancun, Mexico.

[7] Brian S. Butler. 2001. Membership Size, Communication Activity, and Sustainability: A Resource-Based Model of Online Social Structures. *Information Systems Research* 12, 4 (2001), 346–362.

[8] Brian S. Butler and Xiaoqing Wang. 2011. The Cross-Purposes of Cross-Posting: Boundary Reshaping Behavior in Online Discussion Communities. *Information Systems Research* 23, 3-part-2 (Sept. 2011), 993–1010.

[9] Julie Campbell, Cecilia Aragon, Katie Davis, Sarah Evans, Abigail Evans, and David Randall. 2016. Thousands of Positive Reviews: Distributed Mentoring in Online Fan Communities. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. Association for Computing Machinery, New York, NY, USA, 691–704.

[10] Glenn R. Carroll. 1985. Concentration and Specialization: Dynamics of Niche Width in Populations of Organizations. *Amer. J. Sociology* 90, 6 (May 1985), 1262–1283.

[11] Glenn R. Carroll and Anand Swaminathan. 2000. Why the Microbrewery Movement? Organizational Dynamics of Resource Partitioning in the U.S. Brewing Industry. *Amer. J. Sociology* 106, 3 (2000), 715–762.

[12] Eshwar Chandrasekharan, Umashanthi Pavalanathan, Anirudh Srinivasan, Adam Glynn, Jacob Eisenstein, and Eric Gilbert. 2017. You Can't Stay Here: The Efficacy of Reddit's 2015 Ban Examined through Hate Speech. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW (Dec. 2017), 31:1–31:22.

[13] Eshwar Chandrasekharan, Mattia Samory, Shagun Jhaver, Hunter Charvat, Amy Bruckman, Cliff Lampe, Jacob Eisenstein, and Eric Gilbert. 2018. The Internet's Hidden Rules: An Empirical Study of Reddit Norm Violations at Micro, Meso, and Macro Scales. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW (2018), 32:1–32:25.

[14] Shuo Chang, Vikas Kumar, Eric Gilbert, and Loren G. Terveen. 2014. Specialization, Homophily, and Gender in a Social Curation Site: Findings from Pinterest. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*. Association for Computing Machinery, New York, NY, USA, 674–686.

[15] Kathy Charmaz. 2015. *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis* (second ed.). SAGE, Thousand Oaks, California.

[16] Tiago Cunha, David Jurgens, Chenhao Tan, and Daniel Romero. 2019. Are All Successful Communities Alike? Characterizing and Predicting the Success of Online Communities. In *The World Wide Web Conference (WWW '19)*. Association for Computing Machinery, New York, NY, USA, 318–328.

[17] Srayan Datta and Eytan Adar. 2019. Extracting Inter-Community Conflicts in Reddit. *Proceedings of the International AAAI Conference on Web and Social Media* 13 (July 2019), 146–157.

[18] Srayan Datta, Chanda Phelan, and Eytan Adar. 2017. Identifying Misaligned Inter-Group Links and Communities. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (Dec. 2017), 37:1–37:23.

[19] Daniel DellaPosta, Yongren Shi, and Michael Macy. 2015. Why Do Liberals Drink Lattes? *Amer. J. Sociology* 120, 5 (March 2015), 1473–1511.

[20] Shira Dvir-Gvirsman. 2017. Media Audience Homophily: Partisan Websites, Audience Identity and Polarization Processes. *New Media & Society* 19, 7 (July 2017), 1072–1091.

[21] Samer Faraj, Georg von Krogh, Eric Monteiro, and Karim R. Lakhani. 2016. Online Community as Space for Knowledge Flows. *Information Systems Research* 27, 4 (Dec. 2016), 668–684.

[22] Casey Fiesler and Brianna Dym. 2020. Moving Across Lands: Online Platform Migration in Fandom Communities. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (May 2020), 042:1–042:25.

[23] Casey Fiesler, Jialun" Aaron" Jiang, Joshua McCann, Kyle Frye, and Jed R. Brubaker. 2018. Reddit Rules! Characterizing an Ecosystem of Governance.. In *Proceedings of the AAAI International Conference on Web and Social Media*. AAAI, Stanford, CA, 72–81.

[24] Casey Fiesler, Shannon Morrison, R. Benjamin Shapiro, and Amy S. Bruckman. 2017. Growing Their Own: Legitimate Peripheral Participation for Computational Learning in an Online Fandom Community. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 1375–1386.

[25] Jeremy Foote, Darren Gergle, and Aaron Shaw. 2017. Starting Online Communities: Motivations and Goals of Wiki Founders. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 6376–6380.

[26] Brendan J. Frey and Delbert Dueck. 2007. Clustering by Passing Messages Between Data Points. *Science* 315, 5814 (Feb. 2007), 972–976.

[27] Janet Fulk, Andrew J. Flanagin, Michael E. Kalman, Peter R. Monge, and Timothy Ryan. 1996. Connective and Communal Public Goods in Interactive Communication Systems. *Communication Theory* 6, 1 (1996), 60–87.

[28] Mark S. Granovetter. 1973. The Strength of Weak Ties. *Amer. J. Sociology* 78, 6 (May 1973), 1360–1380.

[29] Catherine Grevet, Loren G. Terveen, and Eric Gilbert. 2014. Managing Political Differences in Social Media. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*. Association for Computing Machinery, New York, NY, USA, 1400–1408.

[30] Aaron Halfaker, R. Stuart Geiger, Jonathan T. Morgan, and John Riedl. 2013. The Rise and Decline of an Open Collaboration System: How Wikipedia's Reaction to Popularity Is Causing Its Decline. *American Behavioral Scientist* 57, 5 (May 2013), 664–688.

[31] Jack Hessel, Chenhao Tan, and Lillian Lee. 2016. Science, Askscience, and Badscience: On the Coexistence of Highly Related Communities. In *Tenth International AAAI Conference on Web and Social Media*. 11. arXiv:1612.07487

[32] Benjamin Mako Hill and Aaron Shaw. 2019. Studying Populations of Online Communities. In *The Oxford Handbook of Networked Communication*, Brooke Foucault Welles and Sandra González-Bailón (Eds.). Oxford University Press, Oxford, UK, 173–193.

[33] Serena Hillman, Jason Procyk, and Carman Neustaedter. 2014. 'alksjdf;Lksfd': Tumblr and the Fandom User Experience. In *Proceedings of the 2014 Conference on Designing Interactive Systems (DIS '14)*. Association for Computing Machinery, New York, NY, USA, 775–784.

[34] Itai Himelboim, Kaye D Sweetser, Spencer F Tinkham, Kristen Cameron, Matthew Danelo, and Kate West. 2016. Valence-Based Homophily on Twitter: Network Analysis of Emotions and Political Talk in the 2012 Presidential Election. *New Media & Society* 18, 7 (Aug. 2016), 1382–1400.

[35] Sohyeon Hwang and Jeremy D. Foote. 2021. Why Do People Participate in Small Online Communities?

[36] Thomas J. Johnson, Shannon L. Bichard, and Weiwu Zhang. 2009. Communication Communities or "CyberGhettos?": A Path Analysis Model Examining Factors That Explain Selective Exposure to Blogs. *Journal of Computer-Mediated Communication* 15, 1 (Oct. 2009), 60–82.

[37] Charles Kiene, Jialun "Aaron" Jiang, and Benjamin Mako Hill. 2019. Technological Frames and User Innovation: Exploring Technological Change in Community Moderation Teams. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 44:1–44:23.

[38] Charles Kiene, Andrés Monroy-Hernández, and Benjamin Mako Hill. 2016. Surviving an "Eternal September": How an Online Community Managed a Surge of Newcomers. In *Proceedings of the 2016 ACM Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, 1152–1156.

[39] Robert E. Kraut, Paul Resnick, and Sara Kiesler. 2012. *Building Successful Online Communities: Evidence-Based Social Design*. MIT Press, Cambridge, MA.

[40] Srijan Kumar, William L. Hamilton, Jure Leskovec, and Dan Jurafsky. 2018. Community Interaction and Conflict on the Web. In *Proceedings of the 2018 World Wide Web Conference (WWW '18)*. International World Wide Web Conferences Steering Committee, Lyon, France, 933–943.

[41] Karim R. Lakhani and Eric von Hippel. 2003. How Open Source Software Works: "Free" User-to-User Assistance. *Research Policy* 32, 6 (2003), 923–943.

[42] Cliff Lampe and Paul Resnick. 2004. Slash(Dot) and Burn: Distributed Moderation in a Large Online Conversation Space. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04)*. Association for Computing Machinery, New York, NY, USA, 543–550.

[43] Cliff Lampe, Rick Wash, Alcides Velasquez, and Elif Ozkaya. 2010. Motivations to Participate in Online Communities. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems*. ACM, Atlanta, Georgia, USA, 1927–1936.

[44] Alex Leavitt and John J. Robinson. 2017. The Role of Information Visibility in Network Gatekeeping: Information Aggregation on Reddit during Crisis Events. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. Association for Computing Machinery, Portland, Oregon, USA, 1246–1261.

[45] Jan Marco Leimeister and Helmut Krcmar. 2005. Evaluation of a Systematic Design for a Virtual Patient Community. *Journal of Computer-Mediated Communication* 10, JCMC1041 (July 2005).

[46] Yuyang Liang. 2017. Knowledge Sharing in Online Discussion Threads: What Predicts the Ratings?. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 146–154.

[47] Zhiyuan Lin, Niloufar Salehi, Bowen Yao, Yiqi Chen, and Michael S. Bernstein. 2017. Better When It Was Smaller? Community Content and Behavior after Massive Growth.. In *ICWSM*. AAAI, Montreal, Canada, 132–141.

[48] Eden Litt and Eszter Hargittai. 2016. "Just Cast the Net, and Hopefully the Right Fish Swim into It": Audience Management on Social Network Sites. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. Association for Computing Machinery, New York, NY, USA, 1488–1500.

[49] John Lu, Sumati Sridhar, Ritika Pandey, Mohammad Al Hasan, and George Mohler. 2019. Investigate Transitions into Drug Addiction through Text Mining of Reddit Data. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19)*. Association for Computing Machinery, New York, NY, USA, 2367–2375.

[50] Xiao Ma, Justin Cheng, Shankar Iyer, and Mor Naaman. 2019. When Do People Trust Their Social Groups?. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–12.

[51] A. E. Marwick and danah boyd. 2011. I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience. *New Media & Society* 13, 1 (Feb. 2011), 114–133.

[52] Adrienne Massanari. 2017. #Gamergate and The Fappening: How Reddit's Algorithm, Governance, and Culture Support Toxic Technocultures. *New Media & Society* 19, 3 (March 2017), 329–346.

[53] J. Nathan Matias. 2019. The Civic Labor of Volunteer Moderators Online. *Social Media + Society* 5, 2 (April 2019), 1–12.

[54] Miller McPherson, Lynn Smith-Lovin, and James M Cook. 2001. Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology* 27, 1 (Aug. 2001), 415–444.

[55] Meredith Ringel Morris, Jaime Teevan, and Katrina Panovich. 2010. A Comparison of Information Seeking Using Search Engines and Social Networks. In *Fourth International AAAI Conference on Weblogs and Social Media*.

[56] Meredith Ringel Morris, Jaime Teevan, and Katrina Panovich. 2010. What Do People Ask Their Social Networks, and Why? A Survey Study of Status Message Q&a Behavior. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1739–1748.

[57] Ahlam Muhtaseb and Lawrence R. Frey. 2008. Arab Americans' Motives for Using the Internet as a Functional Media Alternative and Their Perceptions of U.S. Public Opinion. *Journal of Computer-Mediated Communication* 13, 3 (April 2008), 618–657.

[58] Wanda J. Orlikowski. 1992. Learning from Notes: Organizational Issues in Groupware Implementation. In *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work (CSCW '92)*. ACM, New York, NY, USA, 362–369.

[59] Ulrike Pfeil, Panayiotis Zaphiris, and Chee Siang Ang. 2006. Cultural Differences in Collaborative Authoring of Wikipedia. *Journal of Computer-Mediated Communication* 12, 1 (2006), 88–113.

[60] Nathaniel Poor. 2014. Computer Game Modders' Motivations and Sense of Community: A Mixed-Methods Approach. *New Media & Society* 16, 8 (Dec. 2014), 1249–1267.

[61] Catherine M. Ridings and David Gefen. 2004. Virtual Community Attraction: Why People Hang out Online. *Journal of Computer-Mediated Communication* 10, 1 (Nov. 2004).

[62] Martin Ruef, Howard E. Aldrich, and Nancy M. Carter. 2003. The Structure of Founding Teams: Homophily, Strong Ties, and Isolation among U.S. Entrepreneurs. *American Sociological Review* 68, 2 (2003), 195–222.

[63] Joseph Seering, Tony Wang, Jina Yoon, and Geoff Kaufman. 2019. Moderator Engagement and Community Development in the Age of Algorithms. *New Media & Society* (Jan. 2019), 1461444818821316.

[64] Aaron Shaw and Benjamin Mako Hill. 2014. Laboratories of Oligarchy? How the Iron Law Extends to Peer Production. *Journal of Communication* 64, 2 (2014), 215–238.

[65] Kate Starbird. 2012. Crowd Computation: Organizing Information during Mass Disruption Events. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work Companion (CSCW '12)*. ACM, New York, NY, USA, 339–342.

[66] Anand Swaminathan. 2001. Resource Partitioning and the Evolution of Specialist Organizations: The Role of Location and Identity in the U.S. Wine Industry. *Academy of Management Journal* 44, 6 (Dec. 2001), 1169–1185.

[67] Chenhao Tan. 2018. Tracing Community Genealogy: How New Communities Emerge from the Old. In *Proceedings of the Twelfth International Conference on Web and Social Media (ICWSM '18)*. AAAI, Palo Alto, California, 395–404.

[68] Chenhao Tan and Lillian Lee. 2015. All Who Wander: On the Prevalence and Characteristics of Multi-Community Engagement. In *Proceedings of the 24th International Conference on World Wide Web (WWW '15)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 1056–1066.

[69] Nathan TeBlunthuis and Benjamin Mako Hill. 2021. Identifying Competition and Mutualism Between Online Groups. *arXiv:2107.06970 [cs]* (July 2021). arXiv:2107.06970 [cs]

[70] Nathan TeBlunthuis, Aaron Shaw, and Benjamin Mako Hill. 2018. Revisiting "The Rise and Decline" in a Population of Peer Production Projects. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, 355:1–355:7.

[71] Fred Turner. 2005. Where the Counterculture Met the New Economy: The WELL and the Origins of Virtual Community. *Technology and Culture* 46, 3 (2005), 485–512.

[72] Eric von Hippel. 2016. *Free Innovation* (1 edition ed.). The MIT Press, Cambridge, MA.

[73] Xiaoqing Wang, Brian S. Butler, and Yuqing Ren. 2012. The Impact of Membership Overlap on Growth: An Ecological Competition View of Online Groups. *Organization Science* 24, 2 (June 2012), 414–431.

[74] Ryen W. White, Matthew Richardson, and Yandong Liu. 2011. Effects of Community Size and Contact Rate in Synchronous Social Q&a. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 2837–2846.

[75] Xigen Li. 2011. Factors Influencing the Willingness to Contribute Information to Online Communities. *New Media & Society* 13, 2 (March 2011), 279–296.

[76] Justine Zhang, James Pennebaker, Susan Dumais, and Eric Horvitz. 2020. Configuring Audiences: A Case Study of Email Communication. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (May 2020), 062:1–062:26.

[77] Jason Shuo Zhang, Brian Keegan, Qin Lv, and Chenhao Tan. 2021. Understanding the Diverging User Trajectories in Highly-Related Online Communities During the Covid-19 Pandemic. *Proceedings of the International AAAI Conference on Web and Social Media* 5 (2021), 12. arXiv:2006.04816

[78] Xiaoquan (Michael) Zhang and Feng Zhu. 2011. Group Size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia. *American Economic Review* 101, 4 (June 2011), 1601–1615.

[79] Xuan Zhao, Cliff Lampe, and Nicole B. Ellison. 2016. The Social Media Ecology: User Perceptions, Strategies and Challenges. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 89–100.

[80] Haiyi Zhu, Jilin Chen, Tara Matthews, Aditya Pal, Hernan Badenes, and Robert E. Kraut. 2014. Selecting an Effective Niche: An Ecological View of the Success of Online Communities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 301–310.

[81] Haiyi Zhu, Robert E. Kraut, and Aniket Kittur. 2014. The Impact of Membership Overlap on the Survival of Online Communities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 281–290.