

# Communal Quirks and Circlejerks: A Taxonomy of Processes Contributing to Insularity in Online Communities

Kimberley Allison,<sup>1</sup> Kay Bussey<sup>2</sup>

Department of Psychology, Macquarie University,<sup>1,2</sup>  
kimberley.allison@hdr.mq.edu.au,<sup>1</sup> kay.bussey@mq.edu.au<sup>2</sup>

## Abstract

Online communication offers the potential for bridging connections, exposing users to new views and experiences by fostering socially heterogeneous communities. However, in the absence of deliberate attempts to promote diversity, communities may tend towards insularity: a state where members and content are similar or homogenous, and where deviation from these norms is discouraged. This paper presents a taxonomy of processes contributing to insularity, synthesizing findings from a broader longitudinal interview study on engagement with online communities over time with previous literature. Using thematic analysis, sixteen processes were identified which were associated with four broad stages: formation (selective connections, network homophily, shared interests, audience segmentation); propagation (circlejerk, upholding community standards, avoiding conflict, tailoring content); reaction (individual avoidance, collective reaction, mocking deviance, derogating outsiders); and perpetuation (modelling, prior feedback, echo chambers, gatekeeping). These findings highlight the need to consider more diverse mechanisms by which communities become insular, and the role that platform features play in facilitating these processes.

## Introduction

The increasing ubiquity of mediated communication has expanded individuals' ability to connect with others, both within and outside of their existing social circles. These online spaces offer the potential for social capital- benefits to individuals and societies which are derived from social networks and interactions (Ellison, Steinfield and Lampe 2010). Putnam (2000) notes the distinction between bonding capital, which describes the trust and cohesiveness of strong ties within (often homogenous) groups, and bridging capital, where cooperative connections between different (often heterogeneous) groups allow the sharing of information and resources. Each has potential benefits and shortcomings: the stronger and more supportive connec-

tions associated with bonding capital may come at the expense of tolerance of outsiders, while more inclusive bridging connections foster diversity and broaden worldviews but offer less emotional and substantive support (Kobayashi 2010; Putnam 2000; Williams 2006). Both bridging and bonding capital are thought to be necessary for healthy societal functioning (Putnam 2000) and may be developed in online communities (Kobayashi 2010), yet concerns have been raised about a tendency towards insularity in online spaces. Here, we define insularity as a state in which members and content within a community are homogenous or similar, and where deviation from these norms is implicitly or explicitly discouraged; online communities are defined as mediated spaces where users connect with others and/or exchange information or support (Preece 2001).

Bonding capital is inherent in many online communities: platforms based around the articulation of offline networks (e.g. social media) encourage the replication and strengthening of existing connections (boyd and Ellison 2007; Ellison et al. 2010), while those centred on common causes and concerns (e.g. interest-based forums) foster bonding capital by nature of their shared purpose (Williams 2006). Evidence for bridging capital online is more mixed- Ellison and Vitak (2015) posit that social networking sites foster bridging capital as affordances make it easier to maintain connections with weak ties, while other scholars argue that mediated communication by nature facilitates the formation of homogenous, exclusionary communities based on shared identities and interests (*cyber-balkanisation*; Putnam 2000; Sunstein 2007). In more extreme instances these online communities may become echo chambers, reflecting and intensifying shared views, sheltering members from ideological opposition and fostering intolerance (Brainard 2009; Garrett 2009; Hall Jamieson and Cappella 2008). For example, researchers have previously identified these issues in communities characterised by open prejudices (LaViolette and Hogan 2019; Massanari 2017; Mittos et al. 2020) and in online spaces sharing health misinformation (Takaoka 2019). However, these communities do not spring fully-formed into existence, nor is there a clear

delineation between online spaces which are insular and those which are not. Rather, all online spaces exhibit insularity to some degree, ranging from the relatively benign (e.g. keeping forums on-topic) to more extreme forms (e.g. perpetuating toxic cultures). It is therefore important to identify processes that contribute to insularity in online communities.

Previous literature on insularity in online communities is fragmented between disciplines and often limited to the examination of single platforms in isolation. Moreover, insularity is broadly conceptualized as a static state (e.g. echo chambers); where contributing processes (e.g. audience management) are examined, research tends to be framed in terms of other issues (e.g. privacy and disclosure). This paper seeks to highlight the diversity of mechanisms contributing to insularity using data from a longitudinal interview study on individuals' engagement with online communities, drawing from sociological, social psychology and HCI perspectives to inform this work. Here we focus primarily on insularity as it relates to communities as a whole, rather than as it relates to individuals (although we note a certain degree of overlap on platforms where users effectively curate their own communities by choosing who to connect with).

## Related Work

### The Social Psychology of Groups

Individuals derive a core part of their self-concept from the groups they identify with, which shapes both their sense of self and the way they interact with others. Social Identity Theory (Tajfel and Turner 1979) posits that the need for a distinct, positive identity underlies intergroup relations; identifying with a group allows individuals to derive a sense of meaning and belonging, while the behavioural norms that characterise a group provide guidance on how to navigate the world (Fischer and Derham 2016). In particular, categorisation exerts a powerful influence on behaviour in relation to both ingroup and outgroup members, encouraging conformity to group norms, preferential treatment of ingroup members and biases against outgroups (Hogg and Reid 2006). Indeed, the evolution of communities is tightly linked to segregation, homophily and ingroup favouritism (Fu et al. 2012; Masuda and Fu 2015). The impacts of social categorisation and identification have traditionally been investigated with regards to real-world salient groupings (e.g. ethnicity, religion), but have also been replicated in groups where categorisation is explicitly random (Billig and Tajfel 1973). Tendencies to form and adhere to ingroups may have an evolutionary basis, given humans' dependence on their communities for survival (Brewer 1999). Accordingly, such preferences

develop early (Dunham, Baron and Banaji 2008) and are socially valued, with children favouring peers who interact exclusively with the ingroup from pre-school age (Castelli, De Amicis and Sherman 2007).

Social psychology offers some explanation as to how groups distinguish themselves and perpetuate norms. Norms are inferred from the expressed attitudes and behaviours of ingroup members; however while behavioural norms may be relatively accurately identified, discrepancies between individuals' actual and expressed opinions mean that inferred attitudinal norms may be unreliable (Hogg and Reid 2006). In particular, minority or marginalised opinions and experiences may be suppressed in favour of reiterating and affirming normative attitudes: prototypical expressions and shared information are voiced more often in group discussions and considered more valid, while isolation and retaliation against those espousing non-normative positions may further deter the expression of divergent views (the *spiral of silence*; Hogg and Reid 2006; Noelle-Neumann 1974). This reinforces the illusion of consensus within the community, which may in itself perpetuate and polarise opinions (Baron et al. 1996) and bias the behaviour of group members (Sechrist and Stangor 2001). Importantly, the segregation of groups need not be absolute in order to maintain ingroup ties and loyalties; Hewstone (2015) notes that ostensibly diverse communities may resegregate into groups with little productive interaction, while ingroup favouritism has been experimentally demonstrated to increase over time with repeated contact with both the ingroup and the outgroup (Dorrough et al. 2015). Deriving true bridging capital from heterogeneous connections and communities requires productive intergroup contact, ideally involving equal status between groups, common goals, intergroup cooperation and societal support (Pettigrew and Tropp 2006).

### Communicating in Cyberspace

While mediated communication has expanded users' ability to engage with people and perspectives they may not otherwise have been exposed to, technological affordances also allow them to selectively restrict the extent to which they connect, disclose, and hear from others. Consequently, individuals may not be fully benefiting from the potential diversity of experiences in cyberspace. As with offline contexts, successful communication requires both individuals who are willing to disclose, and audiences who are receptive to shared information. Moreover, to truly benefit from the diversity of online communities, users must be willing to disclose and consider information and ideas that are novel, or which may diverge from group norms. These two aspects have largely been considered separately within HCI literature: disclosure through the lens of privacy and boundary regulation, and audience receptiveness through investigations of online group norms and processes.

## Privacy and Boundary Regulation

In all interactions, individuals must negotiate goals and risks of disclosure when sharing information and experiences with others. However, online self-disclosure is further shaped by affordances and features of mediated communication- for example, platforms often merge connections from different segments of an individual's social network into a single homogenous "friend" group by default (*con-text collapse*). Compounded by persistence, visibility and searchability, this increases the risk of content being accessed by unintended and sometimes unwanted viewers (Vitak and Kim 2014). In response, platform users may use behavioural strategies to renegotiate boundaries- for example, across a number of qualitative studies interviewees reported connecting with different audiences on different platforms, segregating audiences or creating multiple accounts within a platform, and employing privacy controls to restrict the audience of shared content (Lampinen et al. 2011; Lampinen, Tamminen and Oulasvirta 2009; Leavitt 2015; Stutzman and Hartzog 2012). These practices are often invoked when sharing content which may receive more mixed reactions (e.g. disclosure of personal or stigmatised experiences, sharing content that diverges from broader societal or platform norms); while this may facilitate disclosure and help-seeking (Andalibi and Forte 2018; Leavitt 2015), this by nature limits the visibility of content that deviates from norms of sharing, allowing standards to be perpetuated.

Self-censorship is also common, as active boundary regulation requires effort and self-efficacy and risks often outweigh motives for self-disclosure (Lampinen et al. 2009, 2011; Vitak and Kim 2014). Importantly, this is not only motivated by privacy concerns- interviewees have also reported avoiding sharing content that may be considered uninteresting, divisive or inappropriate, for fear of social rejection (Vitak and Kim 2014). Marginalised groups may be disproportionately affected by these risks; for example, a study of LGBTQ+ individuals' self-presentation across platforms indicated that risks associated with self-disclosure necessitated restrictions on either the posting or visibility of "high stakes" content (DeVito, Walker, and Birnholz 2018). Self-censorship in response to context collapse may limit the social capital benefits an individual can derive from disclosing to their social networks (Andalibi 2017; Ellison and Vitak 2015), but it also impacts the diversity of content shared as certain types of content (e.g. stigmatised experiences) are reserved for specialised spaces, or withheld entirely.

## Finding your Niche: Segregation and Cyberbalkanisation

The ease with which individuals can create communities and publish content online has allowed the proliferation of spaces catering to specific and niche interests (e.g. Massanari 2017). This allows individuals to form cohesive groups with similar others- for example, those who share political opinions (Sajuria et al. 2015) or LGBTQ+ identi-

ties (Robards et al. forthcoming). Especially for members of marginalised groups, this may represent a safe space for the exploration and expression of aspects of identity that cannot be shared elsewhere (DeVito, Walker et al. 2018; Robards et al. forthcoming). However, research into political discussions online indicates a tendency towards forming echo chambers, with political retweeting largely reflecting partisan segregation (Conover et al. 2011) and commenters on Facebook pages of partisan news organisations preferentially linking to small, largely distinct groups of information sources (Jacobson, Myung and Johnson 2015). Likewise, comments left on blogs reflect a disproportionate level of agreement, although this seems to vary with the genre of posts (Gilbert, Bergstrom and Karahalios 2009).

Similar to offline interactions, individuals are socialised into the norms of a group through exposure and observation, sometimes aided by referral to explicit or formalised guidelines. For example, Facebook users report following behavioural norms inferred from others' activity (Hooper and Kalidas 2012), while new members of the Something Awful forums are expected to have lurked extensively and internalised the site's code of conduct before posting (Pater et al. 2014). Sanctions for deviating from community norms can be even more extreme than in offline contexts, perhaps facilitated by distance and lack of familiarity between users: public humiliation and targeted harassment of those "othered" by the community has been reported on Reddit (Massanari 2017) and Something Awful (Pater et al. 2014), with communities appearing to become less tolerant of transgressors over time (Cheng, Danescu-Niculescu-Mizil and Leskovec 2015). By comparison, the response to transgressions on social media is more individual and passive, often involving the use of platform features (e.g. blocking or hiding) to create distance from deviance without impacting relationships (McLaughlin and Vitak 2011).

Moreover, the spiral of silence appears to replicate online, despite hypotheses that distance and anonymity would weaken the effect. For example, fear of negative reactions discourages individuals from leaving negative online reviews (Askay 2015) and sharing political opinions on social networking sites (Gearhart and Zhang 2015), while conflict avoidance is a commonly reported self-presentation goal on social media (DeVito, Birnholz et al. 2018). In this way, the threat of sanctions encourages behavioural and attitudinal conformity, replicating the standards of the group and encouraging the perspectives of members to align into a collective worldview (Burnett and Bonnici 2003). While these phenomena reflect those observed in offline group interactions, technological affordances facilitate the resegregation of diverse users into homophilic subcommunities and shape the subsequent suppression of non-normative perspectives.

## Current Study

This paper integrates current knowledge from different disciplines and bodies of literature, building a taxonomy of processes contributing to insularity in online communities—that is, behaviours that contribute to the homogeneity of users or content within a community, or discourage diversity or deviation from norms. Here, we present qualitative data from a broader longitudinal study on engagement with online communities, where insularity was identified as a recurring theme. Our research question was as follows:

**RQ:** What processes contribute to insularity within online communities?

## Methods

### Participants

In the interests of gathering a wide range of perspectives and experiences, participants were recruited through online advertisements and word of mouth from three subpopulations: high school students (where engagement with online communities often begins;  $n=12$ , 12-18yrs); tertiary students (where social media use is highest;  $n=18$ , 19-28yrs); and the general population ( $n=14$ , 22-76yrs). The final sample comprised 44 participants (57% female, 43% male), most of whom reported East/South East Asian (45%), Anglo/Celtic (23%) or mixed (16%) ethnic descent. Participants identified themselves as users of 4.7 platforms on average ( $SD=2.0$ , range 1-10); the most common were Facebook (98%), Instagram (64%), Twitter (50%), Snapchat (41%), forums (34%), LinkedIn (27%), Tumblr (23%), gaming communities (21%) and Reddit (20%).

### Procedure

The data reported in this paper were collected in the first phase of a broader longitudinal study on individuals' engagement with online communities over time. Ethical approval for the study was granted by the Macquarie University Human Research Ethics Committee (ref. 5201600539).

After being briefed on the study procedure and indicating consent (including parental consent for those under 18yrs), participants completed individual, semi-structured interviews conducted either in person at a suburban university (73%) or via Skype (27%) by KA. During the interviews, participants were asked about their engagement with online communities, both past and present; this included questions about general patterns of use, norms associated with different communities, and experiences of negative interactions.

Participants were provided with a device to access their online communities for reference, if desired. Interviews lasted 47min on average ( $SD=19$ min, range 14-96min),

and were audio-recorded and transcribed for analysis. All identifying information was removed prior to coding; in reporting findings, participants' names have been replaced with nominated pseudonyms.

### Data Handling and Analysis

All transcripts were coded using a phenomenological approach to thematic analysis, managed in NVivo (version 11). Although participants were not directly asked about processes contributing to insularity during interviews, this was identified as a common theme in responses to other questions (e.g. who participants connected with, what they posted and why, norms of different communities) during data familiarisation and preliminary coding. For example, participants described reported connecting to those known or similar to them and sharing content that adhered to existing norms, and described how community reactions to users and content deviating from these standards discouraged further such behaviour. From here, transcripts were reviewed to aggregate relevant statements, and open inductive coding of all transcripts was completed by KA, supervised by KB. Subsequent axial coding identified patterns within these initial codes, which were then collapsed and organised into broader themes. The reliability of analyses was checked through consistency coding by an independent coder [MC] who was provided with a codebook outlining themes and sub-themes (including illustrative examples), which reached 87% consensus at the sub-theme level.

## Results

Sixteen processes contributing to insularity in online communities were identified from the interview data. These individual- and community-level mechanisms were associated with four broad stages: *formation*, *propagation*, *reaction* and *perpetuation*. Here, we integrate findings from this study with previous literature to more comprehensively synthesise knowledge on processes contributing to insularity.

### Formation

Online communities did not form entirely organically, but were instead carefully chosen and curated by participants through strategic articulation and segregation of connections. Participants often preferentially connected with those similar to them, and further segmented connections across profiles and platforms to counter context collapse.

### Selective Connections

Participants reported being increasingly cautious about who they connected with online, “*only add[ing] certain people*” (Joanna, 25). Many reported not connecting with strangers, consistent with previous research indicating that

these connections often replicate offline networks (boyd and Ellison 2007; Ellison et al. 2010). This was often contrasted with participants' early social media use, which was characterised by more liberal connection practices; over time, participants reflected on and refined their connections in response to privacy violations, unwanted content, and vocational restrictions. For example, one participant reported "*originally connecting with people I knew in real life. But then I decided to make it... people I actually talk to, and who actually are my friends*" (Rebecca, 17). These reports are consistent with the reflections of college students interviewed by Schoenebeck and colleagues (2016), and indicate that the refinement of connection practices occurs at even earlier ages. Some participants also restricted certain groups of people from connecting with them—such as one newly-qualified mental health professional, who decided that "*no-one from my practice shall have me on Facebook*" (Vera, 25). Sometimes restrictions necessitated limiting interactions beyond those specifically prohibited, as anonymity and pseudonymity on some platforms may prevent the identification of other users. One participant, a teacher, reported that:

Because I generally do not know the identity of the people who are posting, I want to play it safe and make sure I'm not... interacting with one of my own students in a way that is not deemed appropriate. (Tim, 31)

Like Zhao and colleagues' (2016) interviewees, participants connected more selectively when adopting new platforms (e.g. Snapchat) to create a more intimate content-sharing environment, in contrast to more widely adopted and connected platforms (e.g. Facebook). Importantly, this process was largely limited to network-based communities which allowed explicit articulation (and restriction) of connections; on other platforms (e.g. YouTube), avoiding connections required avoiding interactions more broadly. While these measures enabled more controlled content sharing, protecting privacy and preventing violations of policy, they also limited users' exposure to new people and perspectives.

### Network Homophily

Individuals often use online communities to connect with others who are similar to them, both demographically and ideologically, mirroring patterns of connections offline (McPherson, Smith-Lovin and Cook 2001). Again, this was particularly the case for network-based communities, where participants described connecting "*with people in my geographic area*" (Jack, 18) or who have "*the same level of intelligence, or... similar backgrounds and education*" (Amelia, 25). In some cases, these offline networks are also similar in their attitudes and outlook; as one noted, "*we have the same thinking... the reason we're friends is cause we have something [in] common*" (Nehal, 18). Another noted that similarities increased through extended contact:

You've got this close social network of people who you interact with frequently... You become very

much like these people, and you have access to the same knowledge and relationships. (Tim, 31)

On larger platforms like Facebook, this trend appeared to be reversing as user-bases expand and diversify; participants reported feeling obliged to accept connection requests from all acquaintances, contributing to context collapse. As previously reported (Zhao et al. 2016), this drove participants to adopt platforms allowing more selective connections, either through design or emergent norms. This included those encouraging replication of specific types of offline connections, for example, professional networks on LinkedIn or educational networks through discussion forums. Other platforms were predominantly used by specific demographic groups— for example, "*WeChat is a very... Chinese-dominated social media platform*" (Ikuy, 23), while "*a lot of people on Tumblr are like... social minorities*" (Becky, 18).

### Shared Interests

Online communities offer platforms for individuals to connect with others who share their interests and passions, which may not be shared by those in their offline networks; indeed, it was often unmet needs to explore these interests which drove participants to these communities. For example, one participant described being "*really into these TV shows*" during high school, "*but no-one at school watched them, so I didn't have anyone to talk to about it. So I would go on the fan forums*" (Amelia, 25); likewise, these spaces may also offer opportunities to connect with others with similar experiences, facilitating disclosure and support-seeking around sensitive or stigmatised topics (Andalibi 2017; Andalibi and Forte 2018). Despite the potential for bridging capital in connecting strangers, here too bonding capital may dominate given the focus on shared causes and concerns (Williams 2006). The proliferation of communities and subcommunities online means that interests which are niche, marginalised or considered deviant can be catered to, with individuals able to "*pick and choose what [they] want to be involved in*" (Ragnar, 23). This was more common on interest- and content-based platforms such as Reddit, where "*everyone who has their own specific likes and stuff can find niches*" (Daniel, 16); Massanari's (2015) ethnographic work indicates this is a major attraction of the platform. However, network-based platforms are increasingly also offering these options (e.g. Facebook groups and pages), and even on platforms where communities are not rigidly defined (e.g. Twitter), participants reported seeking out and connecting with others who shared (content related to) their interests and beliefs. Involvement in interest-based communities tended to follow users' changing interests; as one participant commented, "*once that [interest] kind of dies... my desire to go on Tumblr kind of faded away as well*" (Becky, 18); by contrast, discovering new interests through passive exposure was relatively uncommon in this study. Indeed, Massanari's (2015) work on Reddit suggests the inverse, with users actively resisting platform attempts to expose them to new communities through default subscriptions.

## Audience Segmentation

Similarly, participants reported combatting context collapse by segmenting their connections into subcommunities and restricting certain subgroups from accessing some of their content, both within and between platforms. However, whereas network homophily contributed to homogeneity of *users* within a community, audience segmentation facilitated *content* homogeneity. Dividing their audiences allowed participants to explore aspects of their identity and interests which they felt their broader networks might not understand, like fandom involvement; other researchers have also identified the importance of these spaces for learning about and exploring gender and sexuality amongst LGBTQ+ communities (DeVito, Walker, et al. 2018; Robards et al. forthcoming), and in sharing sensitive or personal information (Andalibi and Forte 2018; Leavitt 2015). Other participants reported presenting a more polished or sanitised persona to avoid disclosing personal information to their more distant (or judgemental) connections, or to avoid conflict. For example, one participant maintained separate private Twitter accounts for friends in Australia and those overseas, “*because I kind of want to keep them distinct and separate*” (Kevin, 13). Another restricted offline connections from accessing her confessional Twitter account, creating “*a safe space [to rant], due to the fact that not many people know who I am*” (Rebecca, 17). For many of these participants, the decision to segment their audiences was prompted by perceived threats to privacy or potential for conflict, rather than previous experiences of conflict.

## Propagation

When sharing content, participants were conscious of how they would fit in with other posts on the platform, and in particular how their content would be received by others. They described constructing posts to maximise their appeal, relevance and consistency with others, and avoided deviating from common standards (including posting content that may cause conflict or controversy). This resulted in rather homogenous posting, with little deviance from norms. When reflecting on the evolution of their content sharing practices, participants here and in previous interviews (Schoenebeck et al. 2016) noted that this became increasingly controlled and curated with time, which they linked to increased maturity and previous negative reactions.

## Circlejerking<sup>1</sup>

Within online communities, individuals often post in ways predicted to appeal to others, reflecting shared communal interests and preferences. Participants reported that some

users deliberately post in ways that will attract more positive responses (e.g. likes, upvotes); one said of Tumblr:

It was like a fun challenge for you to try and get followers [...] Really all I was doing was just posting stuff that I thought that other people would like.  
(Alex, 21)

Similar awareness and pursuit of social validation through peer feedback has been demonstrated amongst adolescent users of Facebook and Instagram (Chua and Chang 2016; Yau and Reich 2019); this study’s findings suggest that these practices extend to other platforms and age groups. This is perhaps unsurprising; social validation is a commonly reported motive for sharing online (Vitak and Kim 2014), and the ability to quantify this (e.g. through likes or karma) appears to make approval more salient and motivate users to consider audience responses when posting (Massanari 2015, 2017; Yau and Reich 2019). In extreme cases, this resulted in shared content becoming similar or repetitive, with users reposting “*low effort*” content (e.g. in-jokes and memes) that would elicit a “*knee-jerk positive reaction*” (Ragnar, 23) from others, rather than more considered, effortful or original posting. Participants felt this was particularly characteristic of platforms like Reddit, which they attributed to the upvoting/downvoting mechanisms used to sort content and (theoretically) indicate quality. Massanari’s (2015) informants similarly suggest that Redditors are highly aware of what is popular within the community and associated karmic rewards, driving them to post in ways that will appeal to others even when this contradicts broader community values around originality.

## Upholding Community Standards

In some online communities, strong norms exist around the content that is posted and shared (e.g. informational, aesthetic). Some participants reported feeling pressured to conform to these norms when posting, while others were deterred from posting by a perceived inability to live up to these standards. This was particularly characteristic of Instagram, where participants reported “*going out of my way to like, manoeuvre things so it looks aesthetically pleasing*” (Rebecca, 17), and “*never [making] any posts because I was like, ‘This isn’t Instagram enough, this isn’t Instaworthy’*” (Becky, 18). However, other participants were similarly deterred from sharing information and opinions in the context of online debates and discussions (e.g. on forums, LinkedIn) as they did not feel they had sufficient knowledge or experience to contribute, relative to others in the community. While here the main reported impact was pressure to conform to standards, perceived positivity biases on social media may also discourage sharing of negative or stigmatised experiences, and inhibiting these disclosures may harm those affected (e.g. by limiting opportunities for support) (Andalibi 2017; Andalibi and Forte 2018).

---

<sup>1</sup> Circlejerking is a slang term referring to the mutual appeal to and gratification of shared interests and tastes within a community

### Avoiding Potential Conflict

Consistent with the spiral of silence theory (Noelle-Neumann 1974), when sharing content online participants were typically cautious about topics considered likely to cause harm, conflict or controversy. They reported considering “*what wouldn’t offend anyone*” (Rebecca, 17), “*is it appropriate- like, will I get any backlash?*” (Bob, 21) and “*if it’s something, you know, that could be taken the wrong way*” (Poppy, 60). Individual differences were evident here; some participants would take care to discuss difficult topics (e.g. politics) sensitively, while others would avoid them altogether. In some cases, this was linked to external regulation of online posting- for example, parental monitoring or legal restrictions. For others these practices were a response to previous negative reactions or conflicts (as with Zhao et al. 2016). In both cases, the result was a tendency to avoid sharing content that deviated from normative views, particularly around more sensitive or polarising topics.

### Tailoring Content to Relevant Audiences

Some participants explicitly considered the relevance of their posts to their audience, with one arguing that “*it’s the duty of any social media to keep content... applicable to people*” (Nikhil, 17). Content thought to relate only to a subset of connections was often posted more privately to this subgroup (e.g. through Facebook groups or Messenger) to avoid “*announcing something that’s meaningless to, you know, whoever*” (Troy, 22); some participants also reported being able to post more freely within these sub-communities, because of their shared interests or purposes.

Other people would just get like annoyed if you’re posting an in-joke photo. There’s no need to post it on public Facebook, you might as well just do it... with the people who are going to actually see it and who know what it is. (Josh, 20)

As such, this process appeared to depend on successful audience segmentation- echoing Zhao and colleagues’ (2016) findings that social media users consider both audiences and community norms when sharing content, and attempt to avoid sharing irrelevant content. While participants often had their audiences’ interests in mind when deciding how and where to share content, these interests were largely inferred from their contacts’ previous activity. Thus one participant’s rule, “*don’t put stuff publicly when it only applies to a certain group of people... [because] not everyone you connect with on Facebook is going to be of the same persuasions*” (Ragnar, 23) may unintentionally restrict community members’ exposure to the more diverse experiences and perspectives of their network, despite ostensible connections.

### Reaction

Individuals responded to content deviating from individual or collective standards and preferences in ways that reduced the likelihood of encountering it again. This occurred at both individual and collective levels, and in-

cluded a range of responses including avoidance, active confrontation or removal of transgressions, and reaffirming shared standards to discourage further deviance within the community.

### Individual Avoidance of Disliked Content

Many participants avoided or removed disliked content posted by others from their social media feeds using functions provided by the platform (e.g. blocking, unfollowing, unfriending). This was most commonly done to avoid minor annoyances or irritations- for example, “*if someone’s constantly posting... especially if it’s like, pointless stuff*” (Nehal, 18), or to protect the individual from graphic, offensive or hurtful content (e.g. online harassment). However, in some cases this was done in response to shared content that conflicted with participants’ views or beliefs:

Any time I see something I don’t like on Facebook, I’ll just hide them. Even if they just express one opinion I don’t agree with, I’ll be like, ‘Oh, bye’. (Alex, 21)

Previous research has suggested these practices are relatively uncommon- for example, only 4% of social media users surveyed by Pew reported blocking, unfriending or hiding others over political disagreements (Rainie and Smith 2012), while participants here applied these sanctions far more liberally, in response to a broader range of perceived violations. This was less common in interest- and content-based communities, perhaps due to the shared interests and purposes of these platforms. Alternatively, they may be linked to the different affordances (e.g. response options, platform structure) of the communities. For example, one participant attempted to avoid forum users who spread negativity, but noted “*you can’t really block someone on a forum. I mean, you can, but [...] you still know they’re there*” (Ernest, 23) as blocked posts disrupted message threads.

### Collective Response to Deviance

Users perceived to go against shared group attitudes, beliefs, interests or values were subject to censure by others in the community. Depending on the platform and community, this could take the form of explicitly attacking transgressors; for example, participants reported retaliation against known hackers and those who sent hate to celebrities:

If somebody’s like, considered an online [aggressor/ threat] you bet they’re going to get a ton of hate [...] cause they don’t want it happening to them. (Kevin, 13)

If [a celebrity] replies to one person [who sent hate], all of the fans will like, message that person and hate on that person. (Wendy, 23)

Elsewhere, perceived transgressions could be as minor as existing as a member of a marginalised group in an online space, as seen in the harassment of women during the #Gamergate movement (Massanari 2017) and beyond (Jane 2014; Stutzman and Hartzog 2012); even aggression

in defence of safe spaces online may become toxic, in some cases deterring those who the practices are intended to protect (Robards et al. forthcoming). Reporting posts that violated community standards was also common; other participants specifically noted Reddit's upvote/downvote system, which automatically hides content that is negatively rated by others. These mechanisms could be used productively to enforce standards of good behaviour, respond to transgressions and remove harmful content; however, many participants noted they are frequently used to censor dissent, as *"if you have a different view, you get like downvoted"* (Daniel, 16). The appropriation of these functions on Reddit was noted in Massanari's (2015) ethnographic work; however, there have been suggestions of similar features being misused to silence prominent feminists on social media both in this study and in the media (Moody 2018).

### **Mocking Deviance**

Individuals who deviate from normative behaviours or views in an online community may be mocked by others. This was more commonly reported in network-based communities where directly confrontation may be less desirable- for example, one participant reported that *"I don't want to get into a fight with someone on the internet- I would see an awful comment, and I would tag a friend in it and be like, 'Look at this dickhead' or something"* (Alex, 21). This seemed to serve dual purposes of reaffirming common values within the ingroup through identifying transgressions, as well as helping those affected to cope. As another participant summarised, *"people just send screenshots and stuff to each other and you just joke about it, and you just get over it"* (Nehal, 18). While this has not been previously evidenced to the same extent in network-based communities, such behaviour reportedly characterises some subversive online spaces where public shaming of deviants reframes community moderation as entertainment, reinforcing norms through ritualised humiliation (Pater et al. 2014).

### **Derogation of Outsiders**

Further ingroups reportedly formed within some online communities, with users enforcing intergroup distinctions through the derogation of those considered to be outsiders or enemies. This largely reflected community membership- for example, one participant noted that in gaming communities, *"communication with opponents is usually restricted to trolling them"* (Daniel, 16), while another noted that some subreddits were derogated or attacked by the broader Reddit community. However, these behaviours also occurred *within* communities in defence of cliques of core users, where *"people on the outskirts are more open to... attacks, or their views not being taken seriously"* (Fiona, 33). More aggressive forms of outgroup derogation have been reported on subversive spaces like 4chan, targeting newcomers who challenge established norms of behaviour;

again, those from marginalised groups (e.g. women, people of colour) are disproportionately affected (Higgin 2013; Manivannan 2013; Marwick 2014). As Marwick (2014) notes, these behaviours systematically other those who do not fit the model of a typical community member, designating them as outsiders and discrediting their potential and actual contributions.

### **Perpetuation**

Once communities have formed and established standards of behaviour, these are perpetuated as new users detect and replicate norms and adjust their behaviour in response to community feedback. In tandem, deliberate and incidental gatekeeping practices in some communities makes intruders and outsiders feel unwelcome, discouraging dissenters from remaining in the group.

### **Modelling**

Communal norms and values were learned by observing the behaviours of others on the platform. For some participants this seemed to be more organic, with standards picked up non-deliberately through *"osmosis from what you see"* (Ernest, 23). For others, it appeared to be a more conscious learning processes of identifying and following trends as *"everyone just wants to be able to connect to other people, so they just... do things that [other] people do"* (Joanna, 25). In both cases, these processes resulted in the replication of existing norms. For example, one participant described how the process of tagging spread through her social network:

My friends started tagging me in things, and I was like, 'Oh, this is kinda cool' [...] it kinda got normalised in the way I use [Facebook], and now I do it. (Becky, 18)

Previous work has identified modelling as an integral part of socialising new members into communities; norms are learned as users identify and replicate posting styles that are favourably received by their peers (e.g. aesthetic styles, political leanings; An et al. 2014; Burnett and Bonnici 2003). Indeed, certain communities (e.g. Something Awful, some subreddits) carry expectations that new users will internalise norms through an extended period of lurking before posting (Massanari 2015; Pater et al. 2014). Here, participants further noted that platform features (e.g. algorithms) which altered the visibility of content based on community re-sponses reinforced the modelling process.

### **Feedback**

Many participants took cues from community responses to their previous posts when sharing content online- for example, one *"learned the social norm that if you always post, you're going to get less likes"* (Vera, 25). They were particularly sensitive to content that was negatively received, and reported that they would avoid sharing this way in future- as one noted, *"if people downvote my post, I'm not going to post similar stuff again"* (Jack, 18). Feed-

back additionally facilitates processes that propagate community standards (i.e. *circlejerking*, *upholding community standards*, *avoiding potential conflict*), as this quantification of audience responses enables individuals to identify and subsequently replicate community standards and indicators of quality. Interestingly, while participants in this study were particularly sensitive to negative responses, previous work with online news communities has indicated that this predicts increased sharing of poor quality posts, while a lack of response is the biggest deterrent of subsequent posting (Cheng, Danescu-Niculescu-Mizil and Leskovec 2014). This divergence may be linked to platform norms (e.g. of connecting with known others on social media) or affordances (e.g. of responses contributing to users' karma on forums) which create reputational implications for posting and responding.

### **Formation of Echo Chambers**

Communities form around collectives with similar viewpoints and interests (particularly on content- and interest-based communities), and content that reflects these shared communal values becomes dominant while ideas that deviate from these norms are discouraged. As one participant noted, "*once you have that community, the community rewards itself, and perpetuates things*" (Ragnar, 23); that is, these processes reflect, intensify and reinforce the culture of the community, echoing concerns of previous researchers (Brainard 2009; Garrett 2009; Hall Jamieson and Cappella 2008). Although this appeared to contribute to community cohesiveness, it came at the expense of fostering diversity of opinion- as another participant highlighted, "*a whole subreddit will be dominated by this one opinion, and there's no way around that... I think it just makes people really narrow minded, to only have one sort of view presented*" (Alex, 21). The absence or visible censure of divergent viewpoints within these communities may further encourage members to converge towards the dominant opinion of the ingroup (Topirceanu et al. 2016), potentially becoming more polarised in their beliefs (Chan and Fu 2017).

### **Gatekeeping and Inaccessibility**

Some communities have conventions (e.g. language, formatting) distinct from those used in broader online communication, signalling ingroup membership and creating feelings of exclusivity- as one participant noted, "*if you understand 4Chan, you feel like you're part of a group... it's very exclusive*" (Wendy, 22). However, another highlighted that even positively- or socially-oriented conventions (e.g. in-jokes) can be "*quite exclusionary to people who are new to the community*" (Ragnar, 23); as Phillips and Milner (2017) note, collective laughter can both build social worlds and exclude outsiders uncomfortable or unfamiliar with the humour shared. Other communities with distinctly toxic cultures may deter some users from contin-

uing to engage- as one participant commented, "*it put a lot of people off. Like they didn't like the toxic environment of the game, so I think a lot of people quit because of that*" (Eric, 16). Notably, this gatekeeping through deterrence can disproportionately affect members of marginalised groups, for whom "edgy" comments and outright hostility communicate prejudice and a lack of safety within the community (Massanari 2015, 2017; Phillips and Milner 2017). Together, this allows communities to remain ostensibly open but functionally closed to outsiders, as entry costs associated with learning conventions of communication or coping with toxicity may be prohibitively high (Burnett and Bonnici 2003; Galston 1999).

## **Discussion**

This paper synthesises findings from a thematic analysis of interview transcripts with previous literature to present a taxonomy of processes contributing to insularity, organised into four broad stages. The *formation* of these communities is largely based on similarities (e.g. demographic or ideological) between group members, with users carefully restricting their audiences and access to shared content to delineate spaces. Communal norms are *propagated* through the posting of content predicted to appeal to others or conform to the standards of the platform, while avoiding content that may cause conflict or diverge from these norms. Individual and communal *reactions* to content that disrupts shared standards discourages further deviance, with ingroups further defended through the mocking or derogation of transgressors and outsiders. Finally, norms are *perpetuated* through modelling and feedback, while indirect gatekeeping communicates exclusivity and discourages intrusion by outsiders. These findings highlight the diversity of mechanisms that may contribute to insularity within online communities; while each in isolation can contribute to homogeneity or similarity of users or content within an online space, they often work in tandem to promote and reinforce adherence to norms and discourage deviance. However, participants' reports emphasise that these processes (and platform features which shape them) are not inherently good or bad, and may be used adaptively to promote positive interactions within communities as well as defensively to maintain the status quo. As Phillips and Milner (2017, p. 124) note in relation to divisive online humour, "*the implications [...] depend entirely on what kind of community it is, and what kind of walls this [...] might strengthen*".

The motivations and considerations of those engaging in these processes appeared to vary with each specific mechanism. Concerns about privacy and boundary regulation shaped participants' decisions about who to connect with (*formation*) and how to share content (*propagation*). While this may facilitate more controlled sharing, particularly

around content that deviates from conventions of posting (e.g. Andalibi and Forte 2018; Leavitt 2015), separating this activity into discrete online spaces reduces others' exposure to these experiences and allows broader norms to be perpetuated. Particularly in network-based communities, this allows bridging connections to be ostensibly maintained without the benefits of mutual exposure to more diverse perspectives. By contrast, social concerns about belonging, validation and cohesion seemed to underlie tendencies to conform to existing community norms of posting (*propagation*; e.g. Vitak and Kim 2014), heightened by awareness of the potential consequences of diverging from these standards (*reaction*; e.g. Noelle-Neumann 1974). Social Identity Theory suggests that these influences may be particularly strong when community membership carries connotations of a shared group identity, where challenging or deviating from these norms may cast transgressors as outsiders (Hogg and Reid 2006; Tajfel and Turner 1979). As a result, these standards may be reinforced and replicated (*perpetuation*), even when this is to the detriment of the community and potential participants (Marwick 2014; Massanari 2015, 2017).

Throughout these processes, the affordances of mediated communication technologies shape the way that individuals and groups navigate the online space. Participants noted that the functions and features offered by certain platforms facilitated the development of homogenous and insular communities. For example, the ability to avoid content without notifying the posters was key to controlling the content users were exposed to, and conversely the ability to manipulate audiences and privacy settings enables individuals to fine-tune disclosure in a manner that is comparable or superior to offline interactions (Vitak and Kim 2014). Of particular note was the use of algorithms to promote content that is predicted to appeal to the tastes of specific individuals and communities, which facilitated processes of feedback and modelling at the expense of exposing users to more diverse content and viewpoints. This was compounded by features which altered post visibility based on community ratings (e.g. Reddit's voting system), which effectively removed non-normative views from the conversation. Here, Rad and colleagues (2018) note a potential conflict at the heart of social media: platforms are motivated to create algorithms which boost content aligned with users' world views as this is known to further drive activity and engagement, but doing so may compromise the potential for bridging capital and undermine the ideological plurality at the heart of functional democracy. Even attempts to deliberately highlight diversity and amplify marginalised voices (e.g. Reddit's default subscriptions) may be thwarted by users' choices (Massanari 2015), perhaps because this may compromise the visibility of valuable and meaningful content (Zhao et al. 2016); moreover, there may be repercussions for minority groups whose presence is highlighted to the broader community (Massanari 2015). At present, true diversity of experiences and perspectives within communities appears to rely on the initiative of users actively seeking to form these bridging connections

(Sajuria et al. 2015), with developing communities appearing to default to insularity in the absence of (or in spite of) intentional efforts to promote diversity.

While this paper provides key exploratory insight into how different processes may work in tandem to enable online communities to become homogenous and insular, the project was limited in scope and methodology. The aim of the broader longitudinal study was to explore engagement across multiple communities and time points, and was consequently underpowered to explore how mechanisms contributing to insularity may vary between individuals and platforms. For example, gaming communities have been argued to foster bridging capital as group members share a common purpose, yet communication often extends beyond the game into other domains and highlights heterogeneity (Cole and Griffiths 2007; Kobayashi 2010). Additionally, while efforts were made to recruit a diverse sample with respect to age, gender and ethnicity, participants were predominantly Australian and their allusions to regional variations in online activity suggest that these processes may vary cross-culturally- for example, in countries with national politically-motivated censorship of online content. To complement this, we have synthesised findings from this study with previous research across disciplines, presenting an integrated taxonomy of processes that may contribute to insularity in online spaces. We envision this work to be useful to those researching toxic, subversive and subcultural communities in understanding how these cultures develop and perpetuate, as well as to platform designers and moderators seeking to promote diversity and inclusivity.

## Conclusions

The ostensibly increasing diversity of online networks may obscure a tendency towards insularity in online communities. While this may not be explicitly intended, drives to maintain privacy, engage with similar others and avoid conflict indirectly contribute to this effect. This study highlights the need to consider broader processes by which online networks become homogenous, the potential for mechanisms to be used protectively and offensively, and the affordances of platform features in facilitating these processes.

## Acknowledgements

The authors thank Michael Child for his assistance with consistency coding interview data.

## References

- An, J.; Quercia, D.; Cha, M.; Gummadi, K.; and Crowcroft, J. 2014. Sharing political news: The balancing act of intimacy and socialization in selective exposure. *EPJ Data Science* 3(1): 12.
- Andalibi, N. 2017. Self-disclosure and response behaviors in socially stigmatised contexts on social media: The case of miscar-

- riage. *Extended Abstracts of the 2017 CHI Conference on Human Factors in Computing Systems*: 248-253.
- Andalibi, N.; and Forte, A. 2018. Responding to sensitive disclosures on social media: A decision-making framework. *ACM Transactions on Human-Computer Interaction* 25(6): article 31.
- Askay, D. A. 2015. Silence in the crowd: The spiral of silence contributing to the positive bias of opinions in an online review system. *New Media and Society* 17(11): 1811-1829.
- Baron, R. S.; Hoppe, S. I.; Kao, C. F.; Brunzman, B.; Linneweh, B.; and Rogers, D. 1996. Social corroboration and opinion extremity. *Journal of Experimental Social Psychology* 32: 537-560.
- Billig, M.; and Tajfel, H. 1973. Social categorization and similarity in intergroup behaviour. *European Journal of Social Psychology* 3(1): 27-52.
- boyd, d.; and Ellison, N. B. 2007. Social network sites: Definition, history and scholarship. *Journal of Computer Mediated Communication* 13: 210-230.
- Brainard, L. A. 2009. Cyber-communities. In H. K. Anheier and S. Toepler eds., *International encyclopedia of civil society* (pp. 587-600). New York, NY: Springer Science & Business Media.
- Brewer, M. B. 1999. The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues* 55(3): 429-444.
- Burnett, G.; and Bonnici, L. 2003. Beyond the FAQ: Explicit and implicit norms in Usenet newsgroups. *Library and Information Science Research* 25(3): 333-351.
- Castelli, L.; De Amicis, L.; and Sherman, S. J. 2007. The loyal member effect: On the preference for ingroup members who engage in exclusive relations with the ingroup. *Developmental Psychology* 43(6): 1347-1359.
- Chan, C.-H.; and Fu, K.-W. 2017. The relationship between cyberbalkanization and opinion polarization: Time-series analysis on Facebook pages and opinion polls during the Hong Kong Occupy movement and the associated debate on political reform. *Journal of Computer Mediated Communication* 22(5): 266-283.
- Cheng, J.; Danescu-Niculescu-Mizil, C.; and Leskovec, J. 2014. How community feedback shapes user behaviour. In *Proceedings of the Eighth International AAAI Conference on Web and Social Media*. Palo Alto, CA: AAAI Press.
- Cheng, J.; Danescu-Niculescu-Mizil, C.; and Leskovec, J. 2015. Antisocial behavior in online discussion communities. In *Proceedings of the Ninth International AAAI Conference on Web and Social Media*. Palo Alto, CA: AAAI Press.
- Cole, H.; and Griffiths, M. D. 2007. Social interactions in massively multiplayer online role-playing gamers. *Cyberpsychology and Behavior* 10(4): 575-583.
- Conover, M. D.; Ratkiewicz, J.; Francisco, M.; Goncalves, B.; Menczer, F.; and Flammini, A. (2011). Political polarization on Twitter. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*. Menlo Park, CA: AAAI.
- DeVito, M. A.; Birnholz, J.; Hancock, J. T.; French, M.; and Liu, S. 2018. How people form folk theories of social media feeds and what it means for how we study self-presentation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. New York, NY: ACM.
- DeVito, M. A.; Walker, A. M.; and Birnholz, J. 2018. "Too gay for Facebook!": Presenting LGBTQ+ identity throughout the personal social media ecosystem. In *Proceedings of the ACM 2018 Conference on Computer Supported Cooperative Work*. New York, NY: ACM.
- Dorrough, A. R.; Glöckner, A.; Hellmann, D. M.; and Ebert, I. 2015. The development of ingroup favoritism in repeated social dilemmas. *Frontiers in Psychology* 6: article 476.
- Dunham, Y.; Baron, A. S.; and Banaji, M. R. 2008. The development of implicit intergroup cognition. *Trends in Cognitive Sciences* 12(7): 248-253.
- Ellison, N. B.; Steinfield, C.; and Lampe, C. 2010. Connection strategies: Social capital implications of Facebook-enabled communication practices. *New Media and Society* 13(6): 873-892.
- Ellison, N. B.; and Vitak, J. 2015. Social network site affordances and their relationship to social capital processes. In S. S. Sundar ed., *The handbook of psychology of communication technology* (pp. 205-227). Boston, MA: Wiley-Blackwell.
- Fischer, R.; and Derham, C. 2016. Is in-group bias culture-dependent? A meta-analysis across 18 societies. *SpringerPlus* 5: 70.
- Fu, F.; Tarnita, C. E.; Christakis, N. A.; Wang, L.; Rand, D. G.; and Nowak, M. A. 2012. Evolution of in-group favoritism. *Scientific Reports* 2: article 460.
- Garrett, R. K. 2009. Echo chambers online? Politically motivated selective exposure among internet users. *Journal of Computer Mediated Communication* 14, 265-285.
- Gearhart, S.; and Zhang, W. 2015. "Was it something I said?" "No, it was something you posted!" A study of the spiral of silence theory in social media contexts. *Cyberpsychology, Behavior, and Social Networking* 18(4): 208-213.
- Gilbert, E.; Bergstrom, T.; and Karahalios, K. 2009. Blogs are echo chambers: Blogs are echo chambers. In *Proceedings of the 42nd Hawaii International Conference on System Sciences*. Washington, DC: IEEE.
- Hall Jamieson, K.; and Cappella, J. N. 2008. *Echo chamber: Rush Limbaugh and the conservative media establishment*. Oxford, UK: Oxford University Press.
- Hewstone, M. 2015. Consequences of diversity for social cohesion and prejudice: The missing dimension of intergroup contact. *Journal of Social Issues* 71(2): 417-438.
- Higgin, T. 2013. b/lack up: What trolls can teach us about race. *The Fibreculture Journal* (22): 133-150.
- Hogg, M. A.; and Reid, S. A. 2006. Social identity, self-categorization, and the communication of group norms. *Communication Theory* 16(1): 7-30.
- Hooper, V.; and Kalidas, T. 2012. Acceptable and unacceptable behaviour on social networking sites: A study of the behavioural norms of youth on Facebook. *The Electronic Journal Information Systems Evaluation* 15(3): 259-268.
- Jacobson, S.; Myung, E.; and Johnson, S. L. 2015. Open media or echo chamber: The use of links in audience discussions on the Facebook pages of partisan news organizations. *Information, Communication and Society* 19(7): 875-891.
- Jane, E. A. 2015. Flaming? What flaming? The pitfalls and potentials of researching online hostility. *Ethics in Information Technology* 17: 65-87.
- Kim, J.-G.; and Lee, K.-S. 2014. Predicting political orientation of news articles based on user behaviour analysis in social network. *IEICE Transactions on Information and Systems* E97D(4): 685-693.
- Kobayashi, T. 2010. Bridging social capital in online communities: Heterogeneity and social tolerance of online game players in Japan. *Human Communication Research* 36(4): 546-569.

- Lampinen, A.; Lehtinen, V.; Lehmuskallio, A.; and Tamminen, S. 2011. We're in it together: Interpersonal management of disclosure in social network services. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. New York, NY: ACM.
- Lampinen, A.; Tamminen, S.; and Oulasvirta, A. 2009. All my people right here, right now: Management of group co-presence on a social networking site. In *Proceedings of the ACM 2009 International Conference on Supporting Group Work*. New York, NY: ACM.
- LaViolette, J.; and Hogan, B. (2019). Using platform signals for distinguishing discourses: The case of men's rights and men's liberation on Reddit. In *Proceedings of the International AAAI Conference on Web and Social Media 13*(1): 323-334.
- Leavitt, A. 2015. "This is a throwaway account": Temporary technical identities and perceptions of anonymity in a massive online community. In *Proceedings of the ACM 2015 Conference on Computer Supported Cooperative Work*. New York, NY: ACM.
- Manivannan, V. 2013. Tits or GTFO: The logics of misogyny on 4chan's random- /b/. *The Fibreculture Journal* (22), 109-132.
- Marwick, A. 2014. Gender, sexuality and social media. In J. Hunzinger and T. Senft (eds.), *The social media handbook*. New York, NY: Routledge.
- Massanari, A. 2015. *Participatory culture, community and play: Learning from Reddit*. New York, NY: Peter Lang Publishing.
- Massanari, A. 2017. #Gamergate and The Fappening: How Reddit's algorithm, governance, and culture support toxic technocultures. *New Media and Society* 19(3): 329-346.
- Masuda, N.; and Fu, F. 2015. Evolutionary models of in-group favoritism. *F1000 Prime Reports* 7: article 27.
- McLaughlin, C.; and Vitak, J. 2011. Norm evolution and violation on Facebook. *New Media and Society* 14(2): 299-315.
- McPherson, M.; Smith-Lovin, L.; and Cook, J. M. 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27: 415-444.
- Mittos, A.; Zannettou, S.; Blackburn, J.; and De Cristofaro, E. Forthcoming. "And we will fight for our race!" A measurement study of genetic testing conversations on Reddit and 4chan. Paper accepted for publication in *Proceedings of the Fourteenth International AAAI Conference on Web and Social Media*.
- Moody, S. 2018. Why is Facebook protecting my abusers? *The Courier Mail*.
- Noelle-Neumann, E. 1974. The spiral of silence: A theory of public opinion. *Journal of Communication* 24(2): 43-51.
- Pater, J.; Nadj, Y.; Mynatt, E. D.; and Bruckman, A. 2014. Just awful enough- The functional dysfunction of the Something Awful forums. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY: ACM.
- Pettigrew, T. F.; and Tropp, L. R. 2006. A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology* 90(5): 751-783.
- Phillips, W.; and Milner, R. A. 2017. *The ambivalent internet: Mischief, oddity and antagonism online*. Cambridge, UK: Polity Press.
- Preece, J. 2001. Sociability and usability in online communities: Determining and measuring success. *Behaviour and Information Technology* 20(5): 347-356.
- Putnam, R. D. 2000. *Bowling alone: The collapse and revival of American community*. New York, NY: Simon & Schuster.
- Rad, A. A.; Jalali, M. S.; and Rahmandad, H. 2018. How exposure to different opinions impacts the life cycle of social media. *Annals of Operations Research* 268: 63-91.
- Rainie, L.; and Smith, A. 2012. *Social networking sites and politics (vol 12)*. Washington, DC: Pew Internet & American Life Project.
- Robards, B.; Byron, P.; Churchill, B.; Hanckel, B.; and Vivienne, S. Forthcoming. Learning and connecting on Tumblr for LGBTIQ+ young people. In A. McCracken, L. Stein, and A. Cho (eds.), *A Tumblr Book: Platform and Cultures*. Iowa City, IA: University of Iowa Press.
- Sajuria, J.; van Heerde-Hudson, J.; Hudson, D.; Dasandi, N.; and Theocharis, Y. 2015. Tweeting alone? An analysis of bridging and bonding social capital in online networks. *American Politics Research* 43(4): 708-738.
- Schoenebeck, S.; Ellison, N. B.; Blackwell, L.; Bayer, J. B.; and Falk, E. B. 2016. Playful backstalking and serious impression management: How young adults reflect on their past identities on Facebook. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work*. New York, NY: ACM
- Sechrist, G. B.; and Stangor, C. 2001. Perceived consensus influences intergroup behavior and stereotype accessibility. *Journal of Personality and Social Psychology* 80(4): 645-654.
- Stutzman, F.; and Hartzog, W. 2012. Boundary regulation in social media. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*. New York, NY: ACM.
- Sunstein, C. R. 2007. *Republic.com 2.0*. Princeton, NJ: Princeton University Press.
- Tajfel, H.; and Turner, J. 1979. An integrative theory of intergroup conflict. In W. G. Austin and S. Worchel eds., *The social psychology of intergroup relations* (pp. 33-47). Monterey, CA: Brooks/Cole.
- Takaoka, A. J. W. 2019. Searching for community and safety: Evaluating common information shared in online ex-vaxxer communities. In G. Meisel (ed.), *Social computing and social media. Communication and social communities. HCII 2019. Lecture notes in computer science, vol. 11579*. Springer, Cham.
- Topirceanu, A.; Udrescu, M.; Vladutiu, M.; and Marculescu, R. 2016. Tolerance-based interaction: A new model targeting opinion formation and diffusion in social networks. *PeerJ Computer Science*, 2: e42.
- Vitak, J.; and Kim, J. 2014. "You can't block people offline": Examining how Facebook's affordances shape the disclosure process. In *Proceedings of the ACM 2014 Conference on Computer Supported Cooperative Work*. New York, NY: ACM.
- Williams, D. 2006. On and off the 'net: Scales for social capital in an online era. *Journal of Computer Mediated Communication* 11(2): 593-628.
- Yau, J. C.; and Reich, S. M. 2019. "It's just a lot of work": Adolescents' self-presentation norms and practices on Facebook and Instagram. *Journal of Research on Adolescence* 29(1): 196-209.
- Zhao, X.; Lampe, C.; and Ellison, N. B. 2016. The social media ecology. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. New York, NY: ACM