



To the Extreme: Exploring the Rise of a Deviant Culture in a Misogynist Digital Community

Yongren Shi¹ , Kevin Kiley² , and Stephanie M. DiPietro¹

Socius: Sociological Research for a Dynamic World
 Volume 10: 1–18
 © The Author(s) 2024
 Article reuse guidelines:
sagepub.com/journals-permissions
 DOI: 10.1177/23780231241272681
srd.sagepub.com



Abstract

Recent instances of lethal mass violence have been linked to digital communities dedicated to misogynist and sexist ideologies. These forums often begin with discussions of more conventional or mainstream ideas, raising the question about the process through which these communities transform from relatively benign to extremist. This article presents a study of the Reddit incel community, active from mid-2016 to its ban in late 2017, which evolved from a self-help forum to a hub for extremist ideologies. We use computational grounded theory to deduce empirical patterns in forum composition, psychological states reflected in language use, and semantic content before refining and testing an interactional process that explains this change: a shift away from drawing on real-world experiences in discussion toward a greater reliance on cognitively simple symbols of group membership. This shift, in turn, leads to more discussions centered on deviant ideology. The results confirm that understanding the dynamics of conversation—specifically, how ideas are interpreted, reinforced, and amplified in recurrent, person-to-person interactions—is crucial for understanding cultural change in digital communities. Implications for sociology of groups, culture, and interactions in digital spaces are discussed.

Keywords

incel subculture, group deviance, conversation analysis, computational grounded theory, culture.

In recent years, digital communities have become critical spaces for the development of deviant cultures and the recruitment and mobilization of individuals toward violent and nonviolent extremist actions (Karell et al. 2023; Smith et al. 2020; Song et al. 2022; Wojcieszak 2009). Many recent events of domestic terrorism¹ were carried out by men identified as members of the incel subculture (short for “involuntary celibate”), which is part of the broader online misogynist network often called the “manosphere” (Farrell et al. 2019; Hanson, Pascoe, and Light 2023; Ribeiro et al. 2021). Several online forums that served as incubators for the development of misogynist ideology have been forced to shut down after conversations on these boards quickly evolved from discussions of self-help and attractiveness to violent, hostile rhetoric toward women and outsiders.

Previous sociological research on online deviance has used qualitative methods to examine misogynistic language use, behaviors, and subjective understandings of forum users once they transition to extreme viewpoints (Glance, Dover, and Zatzkin 2021; Halpin 2022; O’Donnell and Shor 2022; O’Malley, Holt, and Holt 2022; Regehr 2022), leaving open the question as to how these digital communities rapidly move from relatively benign to extremist. At the same time, sociologists specializing in religious and organizational studies have long been interested in the processes by which a community moves toward novel or extreme views, yet the bulk of this scrutiny has been directed toward offline collectives, including religious sects, cults, and utopian

¹University of Iowa, Iowa City, IA, USA

²North Carolina State University, Raleigh, NC, USA

Corresponding Author:

Yongren Shi, Department of Sociology and Criminology, University of Iowa, 401 North Hall, Iowa City, IA 52242, USA.

Email: yongren-shi@uiowa.edu

¹Examples are the 2018 Tallahassee mass shooting (Department of Homeland Security 2022), the 2018 Toronto van attack (Beauchamp 2018), and the 2021 Plymouth shooting (Weaver and Morris 2021).



communities, where interaction dynamics might be different (Kanter 1972; Lofland and Stark 1965; Rochford 1982; Stark and Bainbridge 1980). This article endeavors to unravel this process by addressing two simple but empirically important questions: What quantitative patterns of behavior and culture emerge as communities evolve toward extreme cultures, and how do interactions among community members cultivate a culture of deviance?

The scale and complexity of social interaction in digital spaces presents a challenge for neatly summarizing social behavior. Online worlds often look and feel quite different from in-person interaction, and complex texts, anonymous identification, and myriad real-time interactions can obscure behavioral patterns in ways that make it difficult to translate theories developed offline to the digital world (Edelmann et al. 2020; Nelson 2020). Given these challenges, this article avoids a conventional hypothesis-driven design common in social science research and instead employs computational grounded theory, an approach that combines computational tools and human interpretation to theorize complex processes in voluminous data (Nelson 2020). We approach our question with three steps: pattern detection, hypothesis refinement, and pattern confirmation. This iterative and inductive approach allows the data to inform and refine the research hypotheses as the analysis progresses rather than adhering strictly to a predefined hypothesis and methodology from the outset.

We focus on the *r/incels* forum on Reddit (described in greater detail in the following), a discussion board active from mid-2016, when it began as a relatively benign self-help group, to its ban by Reddit in late 2017 (Hauser 2017). To address the question of how the forum became deviant, we first use a mixture of qualitative and computational analyses, including manual coding, supervised learning, dictionary methods, and topic models, to “see novel patterns in . . . data” (Nelson 2020:3). In this pattern-detection step, we explore who did what with whom and what the psychological and semantic consequences of the interactions were. These analyses inform a concrete “scientific image” (Becker 2008) of how the forum became deviant that we use to hypothesize an interactional process that potentially explains the rise of deviant content. We then test those hypotheses using conversation data.

Group Deviance and Online Conversations

Although often operating at a geographic, temporal, and social scale that vastly exceeds in-person interaction, online communities enabled by platforms like Reddit demonstrate significant similarities in their interactional structure and use of culture to groups and communities developed in offline contexts, such as religious movements, cults, and utopian communes (Kanter 1972; Lofland and Stark 1965; Stark and

Bainbridge 1980). In Reddit’s semi-closed communities, social sanctions and interaction are restricted to subscribed members, which can lead to the development of unique language styles, jargons, values, and strategies as social boundaries that differentiate members from nonmembers, similar to what is observed in offline tight-knit communities. These features foster the development of unique shared cultures within both kinds of communities (Fine 2012). A group culture can engender a sense of belonging and support for members and cultivate a shared understanding of the meanings of particular linguistic expressions, behaviors, and customs to which members can refer and use as the basis for further interaction (Blumer 1986; Fine 1979). Conversely, a shared culture can also result in the exclusion of outsiders and the potential for the reinforcement of extreme or harmful beliefs and ideologies.

But how do such cultures change within online communities? The prevailing models for explaining macro-level culture change are based on interpersonal contagion, where culture is modeled as transferable entities that diffuse through social connections (Axelrod 1997; DellaPosta, Shi, and Macy 2015; Macy et al. 2003; Shi and Shi 2023). Each individual’s opinion on a social issue can be conceived of as a static state (e.g., thumbs up or down or choice on a Likert scale), subject to updates based on received information from their network neighbors. The model includes an observation step in which a focal individual collects information from social connections of varying characteristics and an update step to revise their internal position based on information received from network neighbors.

However, this observe-and-update model overlooks interactional dynamics in which meanings are created, interpreted, and transferred (Goldberg and Stein 2018). For example, in a dinner-table conversation on a contentious social issue, the meanings produced are not only dependent on the participants’ preexisting dispositions but also contingent on how issues are expressed at any moment in a sequentially structured “talk-in-interaction” (Gibson 2008). An utterance can lead to contentious discussions that deviate from agreement or provide an opportunity for mutually committing to shared values, depending on how audience members interpret and react with their own statements. This (dis)continuation of the topics and sentiments in conversational sequences is often simplified as random perturbations that can be absorbed as a residual in formal opinion models (Axelrod 1997; DellaPosta et al. 2015; Macy et al. 2021). While acknowledging the modeling utility of the contagion model, we challenge the realism of its assumptions and propose to directly study conversations as a window to the collective construction of opinions and their actual changes.

Threaded conversations, the form of interactions on many social media platforms like Reddit, resemble informal, in-person conversations where who speaks and what gets said in which order unfold organically (Goffman 1981; Sacks,

Schegloff, and Jefferson 1978). A threaded conversation begins with an original poster who creates a content post that prompts replies from other participants, which then serve as stimuli for subsequent comments (Choi et al. 2015). Each post is recorded as a time-stamped text, and the sequences of all the comments in a threaded conversation can be represented as a spanning tree with no loops or bidirectional edges. Each thread often has a focused topic, and deviations from the topic are often sanctioned informally through votes or removed by moderators. Thus, the content published in an online forum not only spreads interpersonally but is also reinforced between posts in a conversation, or interpost focus of content (DiMaggio et al. 2018). Formulating a set of hypotheses for social dynamics becomes challenging without a deep and descriptive understanding of how culture, sentiment, and behavior change through this interaction structure. Therefore, we formulate the initial hypothesis for cultural change in an online community.

Initial Hypothesis: The presence of deviant content in reply posts and earlier posts is closely related, and this relationship evolves as the community matures.

In the Results section, we further develop the hypothesis. We follow the computational grounded theory framework (Nelson 2020), which comprises the following steps: pattern detection, where computational tools are used to detect important patterns in social behavior relevant to the research question; hypothesis refinement, involving the qualitative interpretation of these patterns and revising the hypotheses; and pattern confirmation, which entails using systematic data analysis to confirm or reject the refined hypotheses.

Reddit and the Incel Community on Reddit

Reddit is one of the largest online information-sharing and discussion websites in the United States.² The Reddit platform is organized as user-created and maintained communities, called “subreddits,” covering a huge variety of topics. Some focus on topics commonly discussed offline, such as news, politics, hobbies, sports, gaming, and science. Others focus on particular identities or advocate marginalized or deviant ideologies rarely found offline, such as r/exMormon, r/Mensrights, r/Incels, and r/FatPeopleHate. Active Reddit

²It ranks as the sixth most visited website, as of March 2023, with an estimated more than 57 million daily active users globally and more than 100,000 active communities (Reddit Press 2023). According to a 2021 survey (Auxier and Anderson 2021), Reddit usage in the United States is more widespread among younger internet users, with 36 percent of respondents between the ages of 18 and 29 and 22 percent ages 30 to 49 reporting ever using Reddit. The platform’s user base in the United States skews male, with approximately 23 percent of adult men accessing the site compared to only 12 percent of women (Statista 2022).

users typically subscribe to, read, and post on multiple subreddits. Subreddits have different rules for whether any registered user of Reddit can post in the subreddit or whether only users who are members of the subreddit can post in it. For most of its history, the r/incels forum that is the focus of our analysis was open to comments from all Reddit users. Reddit adopts a hands-off approach to content moderation, leaving moderation to volunteer subreddit moderators (Gillett and Suzor 2022; Massanari 2017), which, in turn, paves the way for unorthodox ideologies and hate speech to thrive in certain communities.

Within subreddits, registered users submit content such as URLs, videos, images, and texts, which are commented on and voted upward or downward by other registered users. The core feature that Reddit communities use to foster engagement and signal approval and disapproval with community and broader standards is the “karma system,” in which a user gains (or loses) karma points if their comments or posts to a Reddit community are upvoted (or downvoted).

The Reddit community r/incels, activated in mid-2016, initially positioned itself as a supportive space for men struggling with forming romantic or sexual relationships, but over time, it transformed into a forum characterized by the propagation of hateful content and promotion of sexual and physical violence against women.³ It was banned on November 2, 2017, shortly after Reddit’s announcement of a new policy that prohibited “content that encourages, glorifies, incites, or calls for violence or physical harm” (Hauser 2017). The group’s transition from a self-help space to a platform for hate speech highlighted the risks inherent in unmoderated online environments.

The data set used in our analysis was downloaded and parsed from the Pushshift Reddit data set (Baumgartner et al. 2020), which is a large collection of Reddit data that includes posts, comments, and related metadata and is updated regularly. The r/incels data set contains 39,131 unique users and 1,191,797 posts. Over a course of 17 months, the community underwent rapid expansion. During the month leading up to its shutdown, the platform published roughly 136,000 monthly posts, averaging over 4,500 posts per day, excluding bot accounts. The published posts contain considerable amounts of aggressive or violent content, with 1.34 percent of the posts containing the term “kill*,” 0.65 percent containing “rape” or “raping,” and 2.61 percent containing “hate.*”⁴

³People identifying as incels congregated online long before the creation of the r/incels forum, and the full history of the online movement is too long to document here. Different incel forums on the internet have seen varying cultures centered around misogyny or self-help. Although both cultural strains were prevalent in the forum when it was created on Reddit, we focus on the shift between these two strains over time.

⁴An asterisk serves as a wildcard in searches, standing for any number of unknown characters. For example, “kill*” can mean killer, killing, killed, and so on.

Results

Following computational grounded theory (Nelson 2020), we structure the Results section into three stages—pattern detection, hypothesis refinement, and pattern confirmation—to guide readers through the process of uncovering the community dynamics.

Pattern Detection

To better understand the process of change on the forum, three computational tools are used to detect patterns and trends in the conversation data. First, a supervised machine-learning algorithm is used to classify forum users as either incels or outsiders, tracing changes in the forum's composition over time. Second, a dictionary-based approach is used to quantify the psychological content of posts made by users with these different group labels. Third, we analyze how words and phrases cluster together in posts to deliver cultural meanings that resonate with ingroup members. This allows us to trace the change in the use of topics over the community's life span.

Categorization of membership. Reddit is an ecology system that has a variety of interest groups, a feature common to other social media platforms. Because of this, subreddits like r/incels are often visited by Reddit users who are unfamiliar with or do not share the same concerns as the core subreddit users. An initial reading of forum posts suggests that temporary “outsider” visitors were a frequent presence on the forum, so we seek to classify users as either forum insiders (“incels”) or outsiders. To this end, we use a supervised classification approach, coding a subset of members as either incels or forum outsiders and then training a model to classify the remaining users automatically based on their posts.

For the qualitative categorization, we first collect post texts from a random sample of 1,000 unique users (classified by username)⁵ and conduct line-by-line coding of their posted content to ascertain membership status. Insiders are identified by their use of language and commentary that fall within four broad thematic categories: self-professed incel status (e.g., “I am an incel”), adherence to or endorsement of incel ideology (e.g., “Women fall into 3 categories: ladies, prostitutes, and sluts”), expressions of ingroup affiliation (e.g., “I appreciate that you’ve been more sympathetic to us then [*sic*] most of the outsiders that come to this sub”), and contempt or disdain for those perceived to be infiltrating their virtual space (e.g., “You know nothing of our struggle”). Commentary and language posted by those classified as outsiders fall within the following thematic categories:

declarations of outsider status (e.g., “I’m not an incel either”), challenges to or critiques of incel ideology (e.g., “That seems to be what you incels believe”), ridicule and mockery targeting incels (e.g., “I’m by all definitions a chad and I definitely bullied kids like you in school”), and offers of support, advice, and coping strategies (e.g., “I want you guys to know that you’re not alone, and sometimes perspective from the outside isn’t always a bad thing”). Terms unique to incel subculture (e.g., “normie,” “chad,” “pill”) are frequently used by both insiders and outsiders. In the coding process, we refrain from using such terms as markers of group affiliation. In the end, out of the 1,000 users sampled, 559 are identified as incel members, 290 are identified as outsiders, and the remaining 151 users are unable to be classified clearly.

Using the coded data, we devise a model to classify the remaining 38,131 users as either insiders or outsiders. We use the 559 incel members and 290 outsiders we had previously coded to construct a data set, dividing them into a training set with 70 percent of the participants and a test set with the remaining 30 percent. We did not consider indeterminate users as a separate category due to their infrequent posting, weak linguistic signals in the text, and their absence of characteristics that define a distinct category. For example, some people posted only one ambiguous post and left. Excluding indeterminate users during the learning process can enhance prediction accuracy, and we left this group out of the model because it does not constitute a separate identity. In subsequent predictions, we classified all users as either incel insiders or outsiders.

We then apply standard text preprocessing steps to all posts, converting words into stemmed tokens (reduced to their base or root form), generating n-grams (sequences of adjacent words), and applying other steps outlined in Appendix 1. After these steps, we train a lasso regression model with a fivefold cross-validation framework. The final prediction accuracy of the model on the binary outcomes (i.e., whether a user was classified an incel insider or outsider in the qualitative coding) on the test set is 92 percent. We then use the model to classify the remaining users as either insiders or outsiders.

Figure 1 shows a list of the 20 terms most predictive of either incel or outsider statuses in the training data. Terms that predict incel status include those associated with the incel ideology, such as “normi,” “incel,” “cope,” “chad,” and “beta.” Other terms, such as “face,” “imagin,” and “disgust,” strongly linked to how incels expressed their grievances, are also associated with incel identity, as are words such as “we_are,” which suggest a shared group identity. On the other hand, terms that predict outsider status contain less obvious meanings and tend to be interactionally oriented, such as “you_person,” “you_talk,” “of_you,” and “you_guy.” Additionally, some terms often used by outsiders carry connotations that imply alternative ways to tackle the challenges at hand, such as “would_recommend,” “genuine,” and “sorri_that.”

⁵Reddit users are identified only by username, and we cannot link usernames to peoples’ offline identities or other online identities. It is also possible that Reddit users may have multiple usernames, so the true number of forum users is unknown.

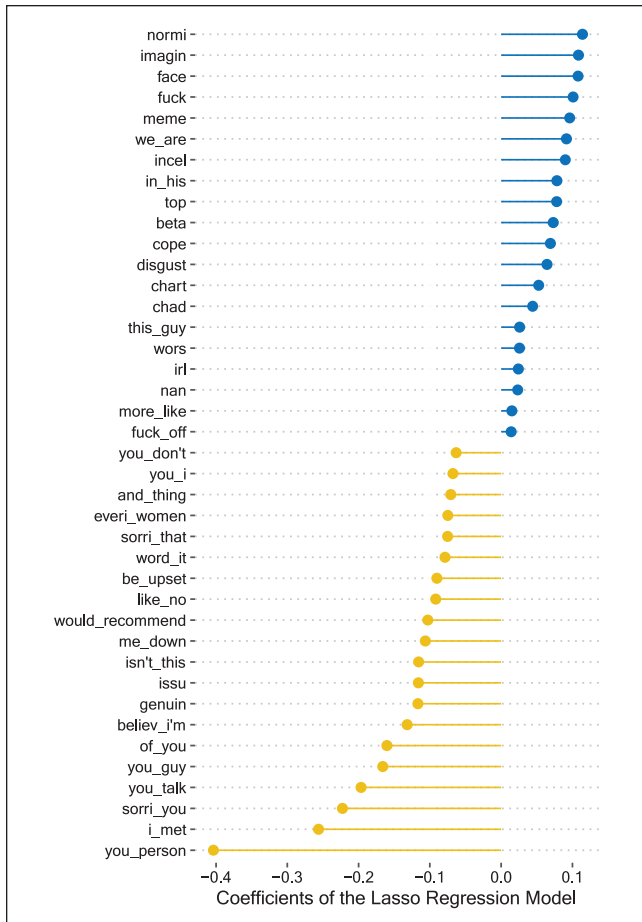


Figure 1. Top 20 terms in the sampled texts most predictive of incel and outsider status.

Note. The blue hue indicates the incel status, and the yellow hue indicates the outsider status.

Figure 2 presents a breakdown of post frequencies and unique user counts categorized by membership status of the poster or user. Initially, the community drew people classified as incel and outsiders at comparable rates, even though people classified as incels posted more often. However, as 2017 progressed, the proportion of people classified as incels within the community increased substantially, the presence of people classified as outsiders diminished, and the amount of content produced by people classified as incels greatly exceeded that produced by outsiders.

Figure 2 highlights the first trend that explains part of the overall shift in aggregate group content—the rising prevalence of people identified as members of the incel community relative to outsiders. However, identifying this trend does not clarify why forum composition changed or, as we note later, why language among the people classified as incels also became more extreme.

Analyzing psychological states. Our preliminary reading of incel and outsider posts suggests three notable characteristics that may signify group identification and deviation: the use of hyperbolic emotional expressions, a strong sense of group identity, and a low level of cognitive engagement, evidenced by the repetitive use of jargon and infrequent engagement in deliberative conversations. This echoes previous work that finds that attachment to online communities has been linked to shifts in members’ social, psychological, and cognitive conditions (Ashokkumar and Pennebaker 2022; Gonzales, Hancock, and Pennebaker 2010).

To explore this shift in the cognitive and emotional content of language, we use a dictionary-based language classifier, Linguistic Inquiry and Word Count (LIWC), which tallies specific terms used in texts that are associated with

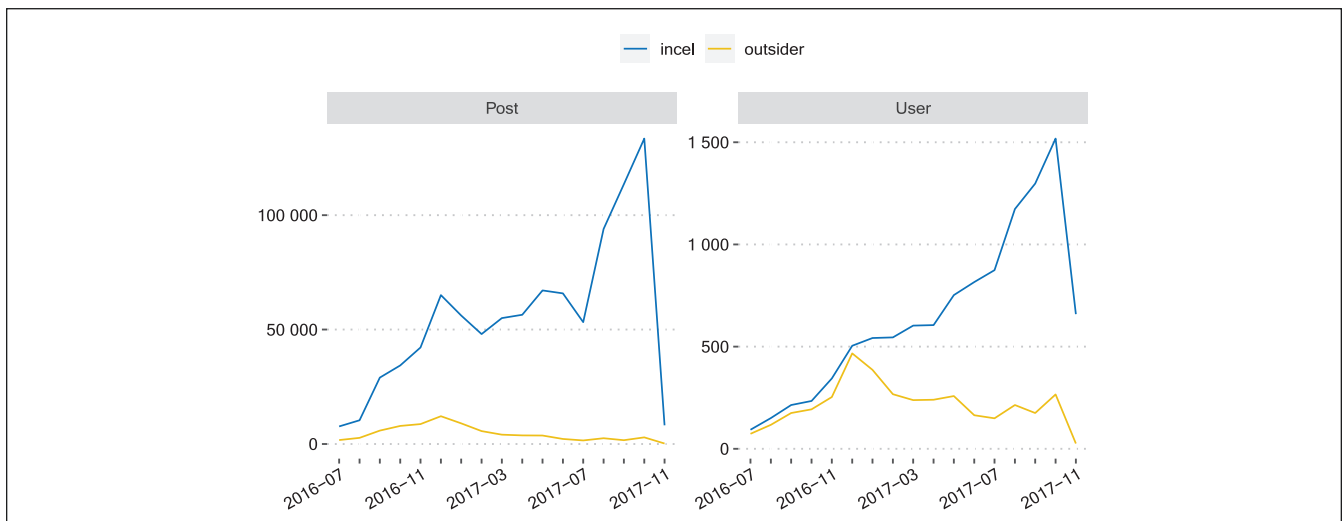


Figure 2. Temporal trends of posts and unique users, broken down by membership status.

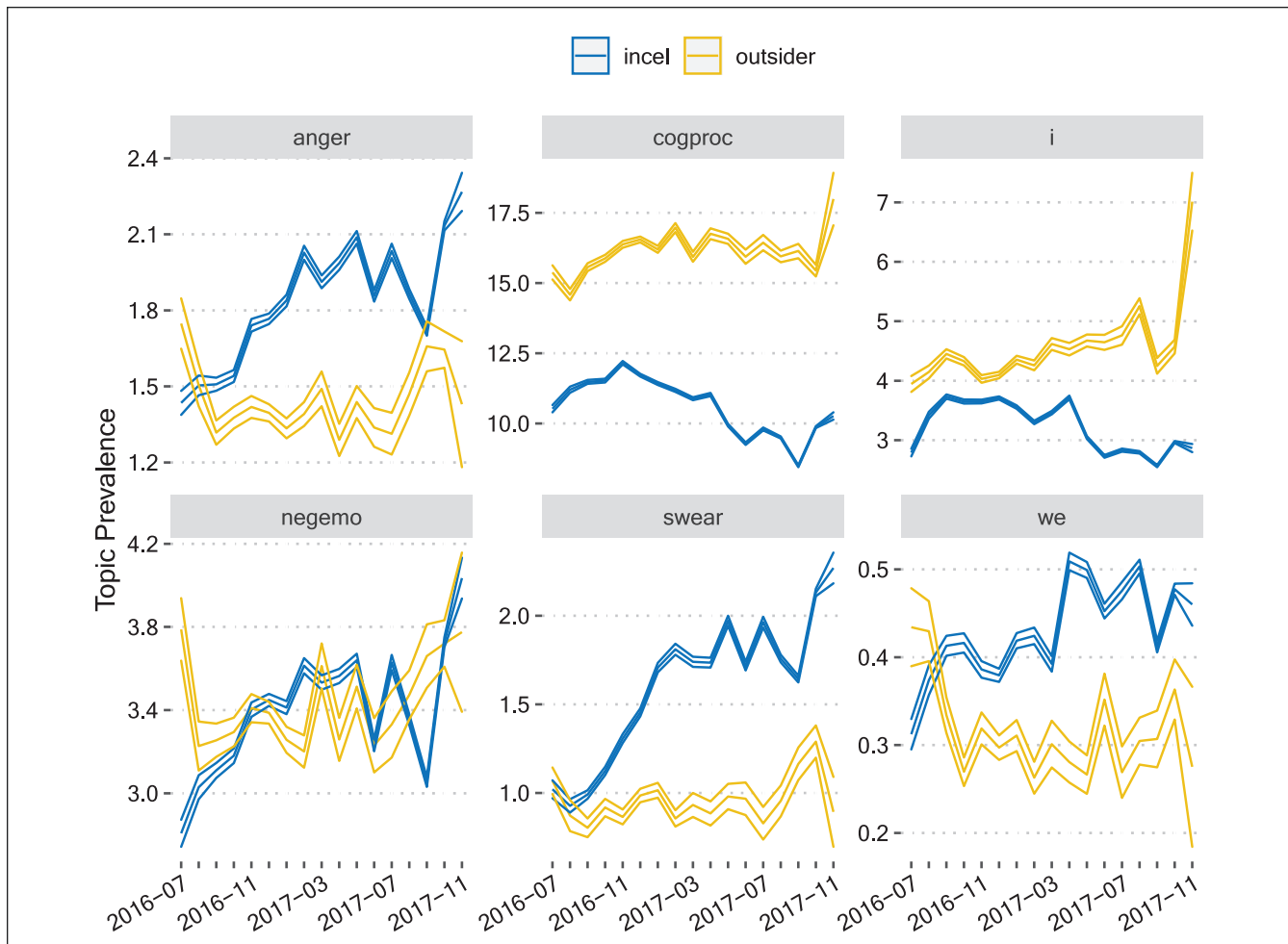


Figure 3. Temporal patterns of social, psychological, and cognitive traits by membership status.

people’s social and psychological states (Pennebaker et al. 2015). LIWC uses multiple dictionaries to quantify the content of text on different dimensions, including the linguistic content (verbs, nouns, quantities), evidence of psychological processes (cognition, affect, social processes), and broad summaries of the psychometric content (“analytic,” “emotional”). LIWC has been used in detecting linguistic patterns in manifestos and communications among lone offenders (Kaati, Shrestha, and Cohen 2016) and radicalization of online communities (Ashokkumar and Pennebaker 2022).

Figure 3 highlights trends in six linguistic dimensions included in LIWC—“anger” (emotional language expressing anger, such as “hate,” “mad,” “angry”), “cogproc” (language indicating cognitive processing, such as “understand,” “think,” “perhaps,” “maybe”), “negemo” (negative emotion words, including “bad,” “hate,” “hurt,” “tired”), “swear” (profanity), and the use of first-person singular (“I”) and plural (“we”) pronouns. We focus on these dimensions because they capture elements that emerged from our qualitative reading.

Consistent with our impressions from preliminary reading, the temporal trends in Figure 3 show a decrease in the cognitive sophistication of language use, increased emotional content, and stronger group identification among users we identify as part of the incel community. First, anger-related language (anger) was consistently more prevalent among the people classified as incels than it was among outsiders, and this prevalence increased over the observed period for users classified as incels, whereas it remained relatively unchanged for outsiders. The incel group’s use of swear words (swear), initially similar to outsiders, quickly rose in the first few months and continued to rise until the end of the forum, whereas the use of profanity by outsiders was low and stable. This disparity suggests a sustained and intensified expression of anger within the incel group, potentially reflective of the community’s underlying frustrations and grievances.

Figure 3 also shows diverging patterns around the use of cognitive processing or questioning language (cogproc) between users classified as incels and those classified

as outsiders. The cognitive processing word list in LIWC dictionary includes terms associated with causation (e.g., “because,” “reason”), self-reflection (e.g., “understand,” “think”), and uncertainty (e.g., “perhaps,” “maybe”; Ashokkumar and Pennebaker 2021, 2022; Hsu et al. 2014). The outsider group exhibits a higher frequency of this kind of language. Initially, 11 percent to 12 percent of words used by incel members were related to cognitive processing, which decreased to 8.8 percent over time.

Finally, users classified as outsiders showed a much higher frequency of first-person singular pronouns (the “I” graph) compared to users classified as incels. Conversely, the use of first-person plurals (“we,” “us,” “our,” “ours,” etc.) was much more common among members identified as incels, and the gaps between incels and outsiders in their use of “I” and “we” grew throughout the forum. This suggests that a stronger sense of community or collective identity and an increasing willingness of individual members to speak on behalf of the group gradually developed among users classified as incels. This finding is in line with recent research on online communities (Ashokkumar and Pennebaker 2022), which noted similar patterns in the use of first-person plural pronouns (e.g., “we,” “us”) to signify collective identity, along with other words related to group affiliation.

These three trends converge to illuminate a key aspect of change in the forum. Over time, people classified as incels decreased their use of language that reflects deep cognitive engagement and analysis and increased their use of first-person plural identification (“we”). This pattern suggests that as a community identity becomes more established, members are less likely to try to convince one another about particular viewpoints. Instead, they increasingly draw on emotional language and profanity. At the same time, this increasing use of first-person plural by group members suggests that members are increasingly willing to speak on behalf of the whole group, potentially suggesting a flattening of apparent differences in worldviews or a more coherent shared worldview within the community over time.

Mapping the semantic content of incel ideology. Although the LIWC dictionaries shed light on the psychological states reflected in how people talk in the forum over time, they cannot illuminate how these changes in psychological states correspond to changes in the content of what is discussed. To explore this, we turn to an unsupervised classification method, topic modeling, to understand differences in language use by people classified as insiders and forum outsiders and how this changed over time. Topic modeling can effectively capture the distributional properties of words and phrases by identifying patterns of co-occurrence across a collection of documents (Arseniev-Koehler et al. 2022; DiMaggio, Nag, and Blei 2013), allowing researchers to extract meaningful insights from the seemingly chaotic flux of conversations (DiMaggio et al. 2018). We use a structural topic

model (Roberts et al. 2013, 2014), an enhanced version of latent Dirichlet allocation, which integrates document-level covariates, such as when something was posted. Structural topic models also allow for topics to be correlated with each other, capturing complex interrelationships among them.

We conduct a number of standard text processing steps before running a 30-topic model. These processing steps and the selection procedure for the number of topics is discussed in the Appendix. The topic keywords, shown in Table 1, are derived from a balanced consideration of high frequency of appearance of the terms under a topic and the extent of exclusivity to that topic (Roberts et al. 2013). Topic labels are assigned based on a qualitative analysis of the top 35 terms in each topic, following an iterative process that alternated between keywords and documents as described in (Nelson 2020). Figure S2 in the Appendix shows a selected sample of posts with their topic distributions. The topic distribution of sample texts suggests that posts typically focus on a small number of topics, with the majority of other topics appearing infrequently or having minimal prevalence.

Our preliminary analysis reveals that users did not employ topics discretely and independently when composing conversations. Instead, certain topics frequently co-occur within the same posts, a phenomenon known as “bundled semantics” (Hanson et al. 2023). To test this, we construct a co-occurrence network to map out its semantic structure based on how frequently topics were used together in posts. Figure 4 shows a network diagram of topics, with connections between topics indicating a high chance that the pairs of topics were used in the same posts. The semantic network shows three distinct topic clusters.

Topics in the first cluster, depicted in red in Figure 4, predominantly revolve around the symbols and metaphors that are central to and unique to the ideology of incels, including terms for stereotypical men (“Chads”) and women (“Stacy”), terms that demarcate the social boundary between incels and the broader society (“normie”), and central aspects of the incel worldview (“pilling”). This cluster also includes generic shared incel experiences, including rejection and denial and dismal treatment, although it is important to note that this topic mostly captures hypotheticals instead of real-world experiences. Topics in this cluster serve as symbols and metaphors for the community, providing novel symbolic frames for the incel adherents and sympathizers to form a new identity (Geertz 1964; Goffman 1986). We consider the topics in this cluster the core incel ideology, consistent with qualitative studies of incel ideology (e.g., Thorburn, Powell, and Chambers 2023). In addition, we posit that these topics comprise the cognitively simple, easily deployed signals of group identity that come to dominate the forum.

The second cluster, colored in yellow in Figure 4, comprises topics that reflect the incels’ focus on physical traits as key determinants of social and romantic success, including facial features, height, weight, and genitalia. Incels

Table 1. 30 Topic Labels and Keywords.

| # | Topic Label | Topic Keywords |
|----|-------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | Society and morality | societi, modern, right, femin, system, moral, power, human, govern, natur, benefit, fault, gene, control, contribut, support, birth, civil, freedom, law |
| 2 | Hate speech | fuck, lol, cunt, retard, bitch, lmao, fuck_fuck, wtf, kek, dumb, stupid, disgust, shut, bro, dude, wow, fuck_hate, damn, fuck_lol, give_fuck |
| 3 | Chad and Stacy | chad, staci, beta, femoid, chad_chad, fuck_chad, roasti, cuck, stacey, beta_provid, dump, chad_fuck, betabux, beta_cuck, orbit, settl, pump, bang, pump_dump, cheat |
| 4 | Love and feeling | love, happi, make_feel, feel, dog, fall, hurt, sad, heart, emot, make_happi, angri, hate, lone, anger, hatr, peopl_feel, affect, song, women_lo |
| 5 | Incel generic | femal, voicel, fellow, male, male_femal, incel, incel_incel, femal_incel, mgtow, standard, inceldom, make_incel, celib, escap, privileg, fellow_incel, lower_standard, misogynist, inferior, involuntari |
| 6 | Virginity | virgin, age, young, marri, lost, shame, loser, gf, past, teen, wait, lose_virgin, late, 20s, forever, kiss, earli, lost_virgin, older, kissless |
| 7 | Dismal treatment | piec, shit, holi, give_shit, holi_shit, treat, piec_shit, give, treat_shit, crap, garbag, peopl_treat, test, fuck_shit, holi_fuck, full, load, respect, peopl_give, ton |
| 8 | Facial features | eye, smile, nose, chin, jaw, light, subhuman, dark, area, open, weak, contact, stare, look, strong, mirror, lip, beauti, close, door |
| 9 | Sex and partner | sex, relationship, prostitut, partner, casual, entitl, intimaci, desir, casual_sex, opposit, romant, sex_relationship, women_sex, escort, sexual, consent, term, sex_drive, hooker, sex_women |
| 10 | Moderation | post, ban, messag, account, mod, pic, thread, user, reddit, delet, send, pm, troll, pictur, op, post_histori, site, match, tinder, check |
| 11 | Mental health | depress, suicid, mental, ill, fuel, mental_ill, life, pain, suffer, kill, therapi, real_life, brain, medic, live_life, attempt, cope, rage, therapist, ldar |
| 12 | Sex attraction | men_women, women_men, men, women_women, women, attract_women, women_attract, attract_men, women_find, hate_women, shallow, attract, women_hate, ugly_men, men_attract, unattract, percent, women_date, men_men, physic_attract |
| 13 | Social life | friend, meet, skill, talk, social_skill, club, social, awkward, parti, make_friend, interact, hang, hobbi, interest, talk_peopl, invit, convers, social_anxieti, lot_peopl, group |
| 14 | Religion | jesus, god, christ, world, earth, follow, hell, real, aw, believ, planet, jesus_christ, real_world, christian, y'all, blackops2cel, version, st, prais, billion |
| 15 | Encouragement | good, bad, luck, good_luck, hope, sound, pretti, work, good_thing, great, hard, thing, feel_good, good_person, job, good_job, pretti_good, gotta, work_hard, idea |
| 16 | Slut shaming | attent, whore, pay, money, valid, she, social_media, compliment, ego, gold, receiv, desper, make_money, worthless, media, seek, slut, paid, pay_attent, call |
| 17 | Wasted effort | time, wast, long, effort, spend, wast_time, spend_time, happen, long_time, heard, reject, worth, multipl, put, period, take, put_effort, time_time, spent, energy |
| 18 | Rejection and denial | dont, im, doesnt, your, care, ur, wont, ive, isnt, tbh, didnt, wouldnt, arent, tho, id, worri, theyr, bother, matter, ppl |
| 19 | Race and ethnicity | white, white_knight, asian, knight, racist, indian, black, race, white_women, white_guy, american, white_girl, trump, white_men, ethnic, countri, minor, asia, black_guy, eastern |
| 20 | Daily routine | day, video, game, play, hour, week, sit, video_game, drink, food, home, watch, smoke, room, dream, wake, minut, movi, hous, night |
| 21 | Normie and pilling | normi, pill, blackpil, advic, black_pill, truth, mock, fail, vent, blue_pill, meme, place, bluepil, pretend, insult, plitud, trigger, come, red_pill, norman |
| 22 | Genitals | dick, suck, vagina, peni, touch, small, porn, stick, size, hand, big, ball, hole, big_deal, wet, tini, bigger, tit, finger, jerk |
| 23 | Height | gt, height, taller, manlet, averag, tall, short, model, shorter, iq, male_model, averag_averag, status, rate, 5'4, face, low, averag_height, lt, ugly_face |
| 24 | Rape and crime | death, rape, autism, sentenc, murder, rapist, victim, deserv, fear, autist, prison, cancer, fire, crime, crimin, violent, die, starv, scare, fight |
| 25 | Beauty and attraction | ugli_guy, nice_guy, hot, girl, guy, guy_girl, nice, cute, ugly_girl, ugly, date, chick, hot_girl, creepi, attract_girl, attract_guy, ugly_peopl, girl_date, girl_guy, creep |
| 26 | Inquiries and responses | question, answer, person, situat, assum, judg, honest, simpl, specif, issu, genuin, chang, problem, perspect, view, curious, attitud, posit, negat, awar |
| 27 | Reason and debate | read, word, sens, argument, make_sens, explain, evid, argu, write, refer, logic, disagre, statement, book, prove, discuss, articl, meant, sourc, wrong |
| 28 | Weight | weight, lift, wear, bodi, muscl, cloth, bald, cut, haircut, makeup, facial, gym, lose_weight, fat, fit, acn, bone, style, shave, structur |
| 29 | School | school, high_school, kid, colleg, year_ago, year, parent, high, ago, mom, dad, mother, class, month, sister, grade, student, father, univers, degre |
| 30 | Plastic surgery | surgeri, save, plastic, suppos, guess, easili, total, plastic_surgeri, turn, think, fine, nope, afford, yeah, poor, doubt, obsess, expens, basic, chanc |

Note. Label color corresponds to the clustered themes, with red being core incel ideology, green being challenges and issues in daily life, yellow being physical characteristics, and gray being uncategorized topics.

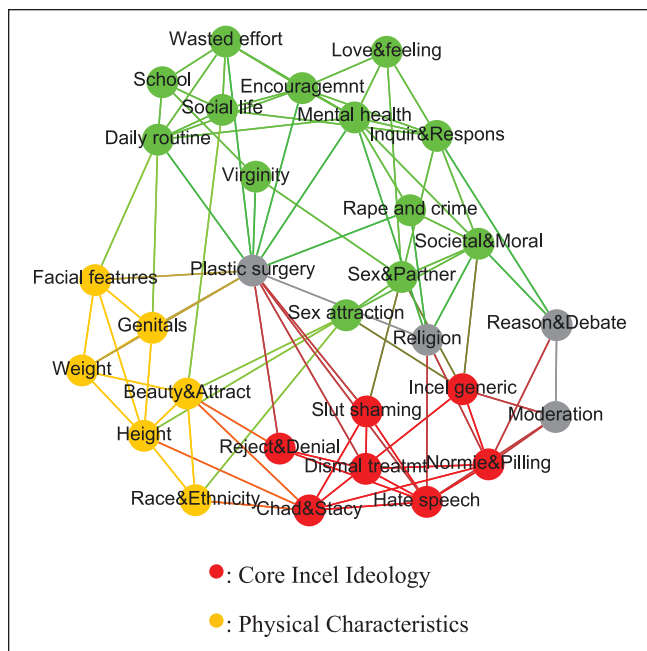


Figure 4. Topic co-occurrence network.
 Note. Nodes represent topics, and edges represent Pearson correlations of topic content in posts between two topics, with those above the threshold of 0.015 being preserved. The color assignment of topics was determined by the network positions and qualitative reading of the topic keywords; red = core incel ideology; yellow = physical characteristics; green = challenges in daily life; gray = unclassified topics.

scrutinize these physical features in fine detail to elaborate perceived attractiveness ideals and express the beliefs that these traits are unfairly distributed and pivotal for relationship success.

The third cluster, what we call “challenges in daily life” and depicted in green in Figure 4, comprises a set of real-world challenges and issues faced by people who identify as incels. This cluster touches on mental health struggles, such as depression, anxiety, and isolation; love and emotion, emphasizing the challenges encountered in forming romantic relationships and coping with unrequited feelings; social life, frequently focusing on social exclusion or difficulties in social interactions; and sex, partners, and incels’ perceived difficulties in finding sexual or romantic partners. The cluster also includes topics reflecting encouragement and positivity, conversations about the investment of time and effort in personal development or attempts at social integration, and daily routines that often reflect a sense of monotony or frustration.

The identified topics within these clusters indicate a thematic emphasis on pseudo-rationalism, which integrates various forms of reasoning, including societal norms, moral frameworks, religious beliefs, and physical appearance ideals. This thematic focus resonates with the scholarly construct of “geek masculinity,” which characterizes masculinity through intellectual pursuits, technical expertise, and a cultural affinity for niche interests such as gaming and science fiction (Ging 2019; Massanari 2017; Van Valkenburgh 2021).

These elements challenge conventional notions of masculinity by emphasizing cognitive abilities and cultural knowledge over traditional markers of male identity, contributing to a nuanced understanding of contemporary gender roles and identities within digital communities. This validates the effectiveness of our topic modeling methodology.

Figure 5 reports the temporal trends of topic proportions of 30 topics over the life span of the community, differentiated by user status (incel or outsider) and theme category. The figure shows that the use of core incel ideology tracks with rising emotional language and declining cognitive processing language highlighted in the previous analysis, with six out of seven topics in the core incel ideology cluster (highlighted in red and presented in the first row of Figure 5) showing an increase in prevalence, with a much higher rate of increase among users classified as incels. For instance, the use of the topic “Chad and Stacy” surged from comprising an average of 1.3 percent of the content of posts produced by incels to 4.7 percent of the content of posts produced by incels. Likewise, incels show a clear increase in their usage of the “Normie and Pilling” topic, with the percentage of their post content containing this language rising from 3.0 percent to 4.5 percent.

The second row in Figure 5 shows that topics related to physical attractiveness also increase in prevalence, with the use of topics such as facial features, height, and racial and ethnic distinctions doubling over the period. The increased prevalence of these topics suggests a reinforcement of incels’ belief of a rigid hierarchy that marginalizes and discriminates against individuals who do not conform to traditional beauty standards. It is noteworthy that the prevalence of these topics in posts made by outsiders also increased. As will be discussed later, the interpost focus of content, rather than ideological deviation of outsiders, may play a crucial role in driving this trend.

The final two rows of Figure 5 show that topics related to challenges in daily life tended to decrease in prevalence over time, especially among users classified as incels. These include topics such as love and feeling (dropping from 4.3 percent to 3.0 percent), sex and partnership (4.1 percent to 2.2 percent), social life (5.5 percent to 3.0 percent), encouragement (5.5 percent to 4.2 percent), and the topics that are related to deliberative language, such as inquiry and response (7.8 percent to 4.0 percent) and reason and debate (7.1 percent to 4.1 percent). The trends among outsiders have also declined, but at a much slower pace.

Although each of these shifts is small in terms of percentage points, aggregating across topics in a particular category clarifies the substantive shift in the kind of content being posted on the forum. In the early months of the forum, 15.9 percent of post content was classified as incel core ideology, compared to 25.2 percent in the final months of the forum. In contrast, about 51.0 percent of content in the early months dealt with the challenges of daily life, which declined to 39.7 percent at the end of the forum.

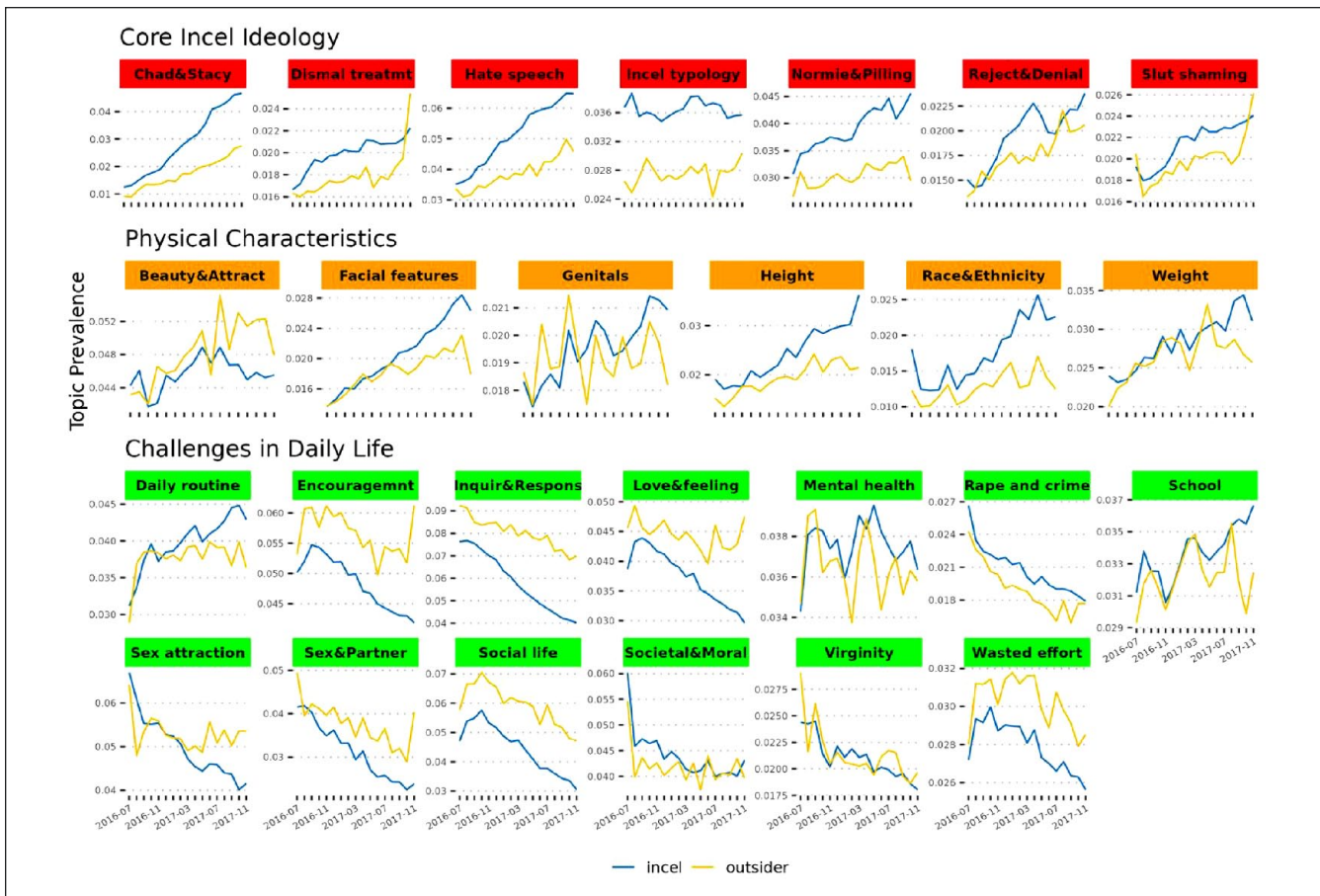


Figure 5. Aggregate change of topics, differentiated by membership status and clustered theme.

Note. Topic prevalence is determined by calculating the average proportion of content attributed to a particular topic.

Hypothesis Refinement

The preceding analyses highlight three co-occurring transitions. First, topics on the forum slowly shifted from discussion of real-world challenges toward use of abstract tokens and core “ideological” considerations, especially among forum insiders. Second, users showed a marked decline in their use of cognitive processing language, suggesting less debate and fewer attempts at persuasion, and an increased reliance on emotion and anger. Third, forum insiders increasingly identified as members of the group (“we”) and became less likely to make statements on their own behalf (“I”). These transitions occurred at roughly similar speeds, with none showing clear temporal precedence, suggesting it is not the case that one caused the others.

We interpret these co-occurring trends as reflecting a mutually reinforcing process. As group members created and agreed on the symbols of the group’s culture (core incel ideology), they used these symbols to both convey group membership and engage with other community members with little cognitive cost. At the same time, discussion of real-world considerations (challenges of daily life) that did not

necessarily conform to the abstract ideology was cognitively costly and potentially divisive. This not only pushed forum insiders to increasingly discuss a purely ideological realm divorced from reality but also increasingly distinguished them from forum outsiders, who neither embraced this ideology nor were conversant in this language. By predominantly discussing the forum’s abstract ideology and deploying it across contexts, incels were able to flatten differences between them, empowering them to speak for the group (“we”). Because the symbolic language provided more of the community’s interpretive and interactive framework, each individual’s need for reflective thought about reality diminished. This relationship between the crystallization of group culture and identities and the simplification of discourse highlights the impact of collective belief systems on the nature of communication within a community.

With these patterns in hand, we step back from analyses to revisit the text data and selectively review samples of conversations from both the early and late stages of the forum. This iterative process provides us with more “on-the-ground” observations, which are essential in transforming the insights into testable hypotheses (Nelson 2020). We select post-reply

pairs where the posts prominently feature core incel ideology, physical characteristics, or challenges of daily life, and then we engage a deep reading of both posts and their replies for the early period and late period of the forum. Specifically, we find all the posts that have a substantial (i.e., about 1 SD above the mean) but similar amount of the content from a topic group (e.g., core incel ideology) during the early period (i.e., first four months) and late period (i.e., last four months) and examine them and their follow-up replies. By doing this, we control the content of the posts drawn from different time periods, allowing a comparative study of the evolution in replying behavior (e.g., Did the replies rely on more deviant content when replying to other deviant posts?). Here, we present three post-reply pairs that satisfy the aforementioned selection condition, which we present in Table 2.⁶ We caution the reader that the following qualitative analysis is not intended as a systematic examination of the data that can derive definitive conclusions but, rather, serves to “either confirm or revise the patterns identified” in the early step (Nelson 2020:26). In the next section, we test the derived hypotheses using quantitative analyses on all eligible post-reply pairs.

The first two rows in Table 2 present posts containing high levels of core incel ideology in the early and late periods and one reply to each. Although both replies address posts of high levels of core incel ideology, they differ significantly. The reply from the early period critically points out the irony in the original poster’s expectation of attention and affection despite their derogatory language toward women, labeling them as “degenerate whores.” Furthermore, this reply advises the original poster to seek therapy. It contains both constructive advice and mainstream reasoning. In contrast, the reply from the late period is marked by a fatalistic and deterministic view of social hierarchies and personal success, reinforcing a bleak and rigid worldview common in incel ideology by suggesting the poster to merely “cope.” At the same time, the language used has become simpler, for example, “ur coping,” which suggests that there is no need to unpack the meaning of such coded words in their communications.

The second two conversations (Table 2, middle rows) highlight a contrast in responses to posts focused on physical characteristics, specifically, regarding height. The reply drawn from the early period challenges the sweeping generalization concluded by the previous poster by pointing out that not all women prefer tall men, offering an alternative view to the incels’ beliefs in physical characteristics in determining romantic success. In contrast, the reply drawn from the late period attempts to downplay the impact of height in dating but emphasizes the importance of facial characteristics. This argument reflects a recurring theme that became prominent in the later stage of incel forum discussions where

physical traits are viewed as the primary determinants of romantic and social success.

The third set of conversations (Table 2, bottom rows) presents responses to posts focused on what we call “real-world challenges.” The reply selected from the early period is compassionate and supportive in response to a post about a sexual assault, and the respondent warns against paying attention to another user’s comment, presumably because it is either harmful or insensitive. The use of “hug” at the end signifies a gesture of empathy and virtual support. In contrast, the reply selected from the late period is simple and stereotypical, characterizing their suffering due to “subhuman looks and subhuman social skills.”

Our qualitative reading suggests a key mechanism at play in the change of the forum over time: interaction and specifically, the evolution of choices (deliberate or not) made by users about when and how to deploy topics in conversation. Specifically, the reading suggests that overall changes in forum content are potentially driven by interpost topic change, or shifts in the content of replies to posts containing the same content. If forum insiders shift how they understand the conversational utility of ideological symbols over time, posts with similar content should produce more responses containing this core incel language as time goes by. This effect would then be compounded by the length of the conversation threads because posts deploying symbols of incel ideology then generate even more posts deploying symbols of incel ideology. Because core incel ideology and physical characteristics are central to the content published in the forum but challenges in life have declined since the beginning, we have the following revised hypotheses:

Hypothesis 1: As time went by, replies were increasingly inclined to reference core incel ideology and physical characteristics in their response to related deviant content in earlier posts.

Hypothesis 2: As time went by, replies were less likely to reference concrete challenges in life in their response to related content in earlier posts.

Should the null hypotheses hold true, that is, no shift in replying behavior over time, it would suggest that the increase in overall deviant content is not a result of conversational dynamics. Instead, it would imply that forum participants have internalized incel ideology, integrating it at similar rates across all types of posts in a thread, be it root, intermediary, or leaf.

Pattern Confirmation: Topic Transitions and Cultural Change

To test the revised hypotheses, we use a simple counting process of post-reply pairs. We define a set of binary indicators for whether a post contains more than the mean level of each topic group (i.e., core incel ideology, physical

⁶The unit of analysis is post-reply pairs, where each post may receive multiple comments and we randomly select only one of these comments for analysis.

Table 2. Sample Conversations from Early and Late Periods.

| Topic Group | Period | Parent Post | Offspring Reply |
|--------------------------|--------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Core incel ideology | Early | I am 10 times the gentleman you stunted little neurotypicals could ever fathom becoming, and yet I get NO PUSSY in this world of sexual chaos and injustice.\n\nThe sheer virtue and impartiality with which I mod this sub entitles me to at least SOME pussy, surely. Yet no female redditor approaches me or acquiesces to my advances... \n\nIt never ceases to confound. | Are you honestly surprised about never being approached when you label us all as degenerate whores simply for existing? You really need some therapy cos it seems you're borderline mentally ill & way over obsessive about getting laid |
| | Late | the bullies at my middle school were the popular kids who became chads and stacies in highschool. they would laugh at me for being weak and defenseless, the worst i had it was getting punched in the face and my glasses breaking (parents couldn't even afford to buy me new glasses for like a month because my prescription was really bad). | ur coping. chad will become a millionaire stock broker while you'll be struggling to make 5 cents from ur self published android app. |
| Physical characteristics | Early | At this point you are just trolling. Its a well known fact that women like tall men and it is so well known in fact that nobody wants to waste resources researching something that obvious. Another reason why many researchers refuse to touch on this topic is because its controversial and its considered sexist to expose women for what they really are. | Sure, many women like tall men, *but not all women*. So, by definition, it isn't a fact. Facts by definition cannot have any exceptions. |
| | Late | Not entirely accurate. Pint-sized cunts hate short men the most. They want big guys that make them feel "protected" | If you go on the short sub-reddit in threads where people post pics you will quickly notice that all the ones with cute GFs are good looking. the ones that are incel and blame it on their height have unflattering faces. It's pretty obvious to me that face is almost everything. Sure being short is a deal breaker to some girls. but even if it was a deal breaker to say 90% of girls (and it's much less than that), leaving 10% of girls being attracted to you... you would have an easy time getting a GF |
| Challenges of daily life | Early | I was sexually assaulted while on a third date with a man that I thought was attractive and successful and kind. Now the idea of him disgusts me and I struggled with intimacy with my current boyfriend because of it. His appearance did not make the situation any less scary and awful. | Please don't listen to, or even skim, killedforbeingincel's comment. It is not worth the time. I hope you feel better soon and are getting help. *hug* |
| | Late | I just want random ass strangers to quit bullying me for how I look. Oh, and no I'm not still in school. Because newflash, the bullying never stops because people never grow up. But that's all cope, that'll never happen. Only time it'll happen is when I blow my brains out. | Are you neurotypical? I feel that most of my inceldom results from both subhuman looks and subhuman social skills. |

Note. The parent posts contain a substantial amount of content from respective topic groups.

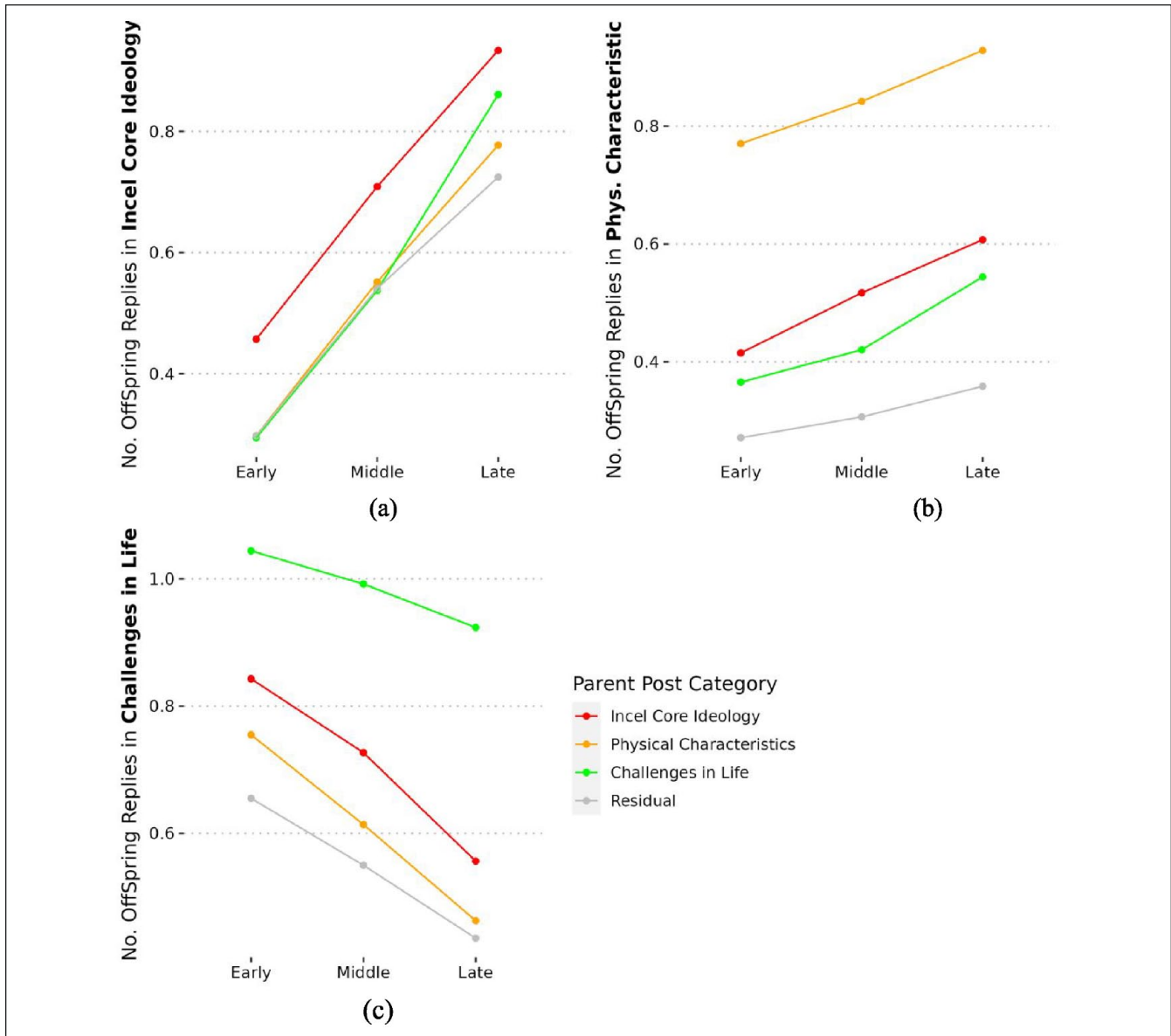


Figure 6. Topic transitions between parent posts and offspring replies during early, middle, and late periods of the community. The values on the vertical axes indicate the numbers of offspring replies in: (a) Incel Core Ideology, (b) Physical Characteristics, and (c) Challenges in Life.

characteristics, or challenges of daily life). For example, a post is considered as focusing on incel core ideology if it includes this type of content more frequently than the average occurrence in all posts.⁷ We then count the frequencies of content categories of the parent posts and offspring replies. For simplicity, we do not differentiate which level that a parent post is located in a conversation tree (e.g., a “parent” post could be an “offspring” reply in a different pair).

⁷When we varied the threshold from the mean to mean plus 1 SD, the overall pattern of results remained consistent.

Figure 6 displays the relationship between parent posts with a specific content category and the number of offspring replies they generate, which may contain either the same or a different content category. Each of the three panels illustrates the numbers of offspring replies related to incel core ideology, physical characteristics, and challenges in life (differentiated by color). Within each panel, we also differentiated the time periods. We chose three periods: The early period ranges from the forum’s inception in July 2016 to November 2016, the middle period extends from February 2017 to May 2017, and the late period begins in August 2017 and continues until the ban in November 2017. For example, in Figure 6a, the

red dot in the early period means that a post in the early period related to incel core ideology would be expected to generate .46 replies containing incel core ideology.

Two interesting observations are evident from this analysis. First, parent posts are most likely to elicit offspring replies containing the same content category. For example, Figure 6a shows that parent posts containing incel core ideology (red line) are significantly more likely to elicit responses with the same incel core ideology compared to a post with different content (yellow, green, and gray lines). The pattern of ordering observed in Figure 6a is also present in the other two panels. However, the gaps between the top line (same content) and the other lines (different content) are much larger in Figures 6b and 6c. This suggests that discussions about physical characteristics and daily life challenges in the community are more likely to stay focused on the same topic than posts about incel ideology.

Second, the three panels of Figure 6 reveal distinct temporal patterns for the different offspring topics. In Figures 6a and 6b, we observe an increase over time in the average number of offspring replies containing incel core ideology and physical characteristics regardless of the parent post content, meaning that posts, regardless of content, were more likely to provoke responses with deviant content in the late period than in the early period. This result supports Hypothesis 1 and suggests a shift among participants toward a more extreme mindset.

Interestingly, Figure 6c shows the opposite trend, where posts of all topics in the later period produce fewer responses that reflect actual real-world experiences. This result supports Hypothesis 2. Also evident from this panel is that the rates of decline in engagement with parent posts about incel core ideology (red) and physical characteristics (yellow) are faster than those concerning daily life challenges (green). This indicates that particularly in the later stages of the community, participants were less likely to return the conversation to real-world examples when the parent posts were related to incel ideology than they were earlier in the forum's life.

The findings support our interpretation that changes in posting behavior and the rise in deviant content reflect the interactional dynamics of the forum. Importantly, although the absolute change in offspring replies is often small, due to the tree-like structure of the conversation threads, these small changes in posting behavior can have a compounding effect over the successive levels of the tree. For instance, in the early stage of the forum, a post containing incel core ideology might result in $0.79 (= 0.47 + 0.47^2 + 0.47^3)$ subsequent posts over three steps, whereas in the later stage, the same type of post could lead to $2.7 (= 0.95 + 0.95^2 + 0.95^3)$ subsequent posts in three successive levels.

Discussion

In recent years, digital communities have become hotbeds for the emergence and propagation of sexist and misogynist hate speech, creating toxic environments for women and

other marginalized groups. The normalization of misogynistic attitudes has also led to several physically violent events (for a list, see Anti-Defamation League 2023). Despite a rich qualitative literature describing the distinctive jargons and ideology that emerge from online forums and despite an exploding interest in studying incel culture and manosphere after the ban of Reddit's r/incels subreddit, there lacks a basic understanding of the group and interactional processes that engender group deviance within semi-closed communities. This is because of a twofold challenge: On the one hand, detailed behavioral trace data are seldom gathered for deviant communities in sociology; on the other hand, due to the fact that clear behavioral patterns are often obscured within myriad complex texts and interactions, it becomes exceedingly challenging to anticipate a ready theoretical framework that would adequately describe change in outcome, a common obstacle facing many computational social science applications (Nelson 2020).

To this end, we leveraged a complete history of conversations from the r/incel community from its inception until its eventual shutdown by the Reddit platform to begin to theorize the process that leads to deviance within an online community. Through a series of computational explorations (Nelson 2020; e.g., machine-learning classification of users, dictionary-based classification of language use, inductive semantic classification), we identified a set of patterns that converged to clarify a theoretical model of deviance in the forum.

By tracking the evolution of psychological states and deviant content, we found that the rise of deviant ideology in the community was associated with the use of cognitively simple language. This mirrors the concept of "newspeak" in George Orwell's 1984, where language is deliberately simplified to restrict thought and control ideology (Orwell 2021). In contrast to Orwell's dystopia, where simplified language is imposed from the top down, the reduction in linguistic complexity within the r/incels community arose organically from interactive dynamics among incel users, fostering a communication style that effectively marginalizes complex, contrary viewpoints. A phenomenon analogous to reduced cognitive complexity is observed in thematic bundling (Hanson et al. 2023), where users frequently combine certain topics in conversations, and these bundled topics exhibit similar trends in rates of change. The growing prevalence of topic bundles, such as core incel ideology and physical characteristics, in conversations reflects a tendency among users to engage in narrow discourses. This facilitates interactions among ingroup members who share a common understanding of thematic content and discourages conversations that deviate from established patterns.

These strands led us to develop and test a theoretical proposition: The character of the forum changed because forum insiders increased the rate at which they deployed abstract symbolic language, shifting discussion away from consideration of concrete daily challenges toward increased discussion of an ideology that was often divorced from real-world

experience. We tested two hypotheses derived from this theoretical model by tracing the interactional process over time, finding that forum users increasingly deployed these symbols of group membership and ideology regardless of the topic being discussed and decreasingly deployed evidence from everyday life to further conversation. These shifts compounded to radically alter the content of the forum over time.

Because these symbols and thematic bundles lack complexity and nuance, they are ill-suited for addressing the messy reality of the problems that people who identify as incels (or who might consider identifying as incels) might encounter. Similarly, because of this mismatch between ideology and reality, discussion of real-world problems poses the possibility of fostering dissensus among group members, a problematic prospect for people who already identify as social outsiders struggling to find belonging. As a result, incels decreased their engagement around real-world issues, what we classified as “challenges in daily life,” and increasingly engaged only around ideas, what we labeled “core incel ideology” and “physical characteristics.” This means the forum was increasingly dominated by discussion of ideas that are rarely tested on a complex reality that might challenge them. This allows users to increasingly identify as members of a community that they perceived to share a similar understanding of the world, akin to the Berger’s (1990) notion of “plausibility structure.”

Our attention here has been on theorizing the process that leads to group deviance, and we do not purport to fully explain every aspect of the forum’s transition. For example, because we cannot directly survey or interview forum participants, our analysis cannot touch on whether some or all users purposefully took these steps or whether they reflected an unconscious feedback process. Similarly, although we explored how the forum transitioned over time, it is not ultimately clear why this forum transitioned in the way it did.⁸

Another important factor that we have not considered in influencing group deviance is moderation. The change of overall culture may occur due to a shift in the community’s rules or changing behavior of moderators whereby certain topics become either encouraged or discouraged. As a result, deviant users may be encouraged to publish significantly more posts that promote extremist views, or outsiders’ views may be removed from the forum. Gillett and Suzor (2022) studied the role of automated moderation (i.e., using bots to remove comments by searching keywords) and active moderation in shaping forum discourse in r/incels and found that content moderation played a significant role in safeguarding the forum’s discourse from external influences. Our analysis, depicted in Figure S3 (SI Appendix), also shows an increase in mentions of a key moderator’s name by incel members.

⁸For example, the transition could be a result of membership replacement as people who could not tolerate the toxic space left and those who shared the same extremist views joined the community. Future research is needed to explore alternative mechanisms.

This moderator enforced strict moderation on posts that challenged or disapproved of prevailing views in the forum. This pattern loosely coincides with the rise of incel ideology in the forum, indicating a potential influence on its growth. However, because the content of the moderated posts was removed from the data set, it is difficult for us to assess their impact, direct or indirect, in a causal manner. We suggest that the bottom-up conversational dynamics may greatly amplify the top-down moderation. Future studies should consider collecting real-time data or data through the wayback machine (Gillett and Suzor 2022), including moderation actions, to directly observe which types of content are likely to be removed and study how they impact conversations.

That being said, the insights developed here potentially inform future research on the formation of communities displaying alternative ideologies. Our results suggest that these ideologies or worldviews potentially emerge as a way to foster a group identity while avoiding challenging aspects of reality. On the one hand, these symbols and metaphors are the fundamental tenets of an ideology that, in the words of Clifford Geertz, were employed to “render otherwise incomprehensible social situations meaningful, to so construe them as to make it possible to act purposefully within them” (Geertz 1964:220). They served as a foundation for a significant social identity that unites incels who may come from disparate backgrounds and possess diverse life experiences. On the other hand, these symbols acted as impediments that alienated outsiders from fully comprehending the challenges confronted by the incel community, solidifying the cultural boundaries. Frequent use of these symbols might have made it more difficult for forum outsiders to positively engage with forum insiders. As these symbols became more prevalent in the forum, it also became more recognizable that someone was not embracing them, making it easier to dismiss these people as “outsiders” who do not understand incels. Although it is beyond our current study to explore whether forum users were aware of the fact that they engage in this process, we encourage future researchers to consider this dimension of cultural complexity—and its connection with group identity—in studies of group deviance.

Finally, our findings on parent-offspring interpost focus and the gradual increase in topic transition in conversations might inform the development of cognitive models for conversational interactions. Unlike person-to-object interactions or indirect observation (Lizardo and Strand 2010), direct conversational interactions are a common yet underexplored aspect of socialization in the sociology of culture. Conversation analysis typically concentrates on the behaviors and strategies of social interactions, such as turn-taking, interruptions, and backchanneling (Gibson 2008; Goffman 1981; Sacks 1992; Sacks et al. 1978). These analyses often overlook the content and cultural significance of conversation topics. As Blumer (1986) conveyed, social interactions are a dual process: They involve signaling to others how to act and interpreting others’ signals. The actions and expressed

opinions in these interactions are not just reflections of pre-existing cognitive dispositions; they are also influenced by the immediate context. Future studies are necessary to explore the complex relationship among conversational dynamics, individual cognitive structures, and broader cultural change within a community.

At the interpersonal level, our analysis of conversational dynamics offers insights into diffusion and cultural change as an alternative mechanism to the prevailing social contagion model, which typically emphasizes explicit social relationships as conduits for diffusion. Due to the anonymity prevalent in internet forums, the conventional contagion model is ineffective for modeling interpersonal dynamics in these contexts. Individuals are less inclined to express privately held but socially unpopular opinions, such as misogynistic views, in their offline social networks, where conflicts in familial and friendship relationships could lead to push-back or strained relationships. However, on anonymous internet forums, individuals have less personal investment in expressing unpopular opinions, and thus they are more susceptible to influence and to influencing others toward the development of deviant ideologies through conversations.

Conclusion

This study focuses on a digital community that emerged as a hub for propagating extremist misogynist ideology across other forums until its eventual shutdown. Taking it as a case study, we found direct evidence of deviation by quantifying and analyzing the progression of deviant content published in the community in its life span. Our research offers critical insights for sociology, highlighting how online platforms, even without the structural features often associated with the development of alternative ideologies (physical and social separation, charismatic leaders, and organizational hierarchies), can influence deviance and group dynamics. The observed trend of thematic consistency in parent-offspring post interactions underscores the role of digital spaces in reinforcing specific ideologies, contributing to the formation of echo chambers. This pattern challenges traditional sociological theories on group dynamics and emphasizes the need to consider the unique influences of conversational interactions on collective ideologies and behaviors. This research underscores the importance of integrating the dynamics of digital spaces into sociological theories to better understand the evolution of group ideologies and behaviors in contemporary society.



Acknowledgments

We gratefully acknowledge the valuable feedback to the article provided by Freda Lynn, Steve Hitlin, Regan Smock, and the participants at the Public Policy Center, the University of Iowa, 2022 American Sociological Association Annual Meeting, and 2022 International Conference for Computational Social Science.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: We are grateful of the funding support from the National Science Foundation (No. 2048670) and the Office of the Vice President for Research at the University of Iowa.

ORCID iDs

Yongren Shi  <https://orcid.org/0000-0002-1876-6237>
Kevin Kiley  <https://orcid.org/0000-0002-7613-2814>

Supplemental Material

Supplemental material for this article is available online.

References

- Anti-Defamation League. 2023. "Incels (Involuntary Celibates)." <https://www.adl.org/resources/background/ncels-involuntary-celibates>.
- Arseniev-Koehler, Alina, Susan D. Cochran, Vickie M. Mays, Kai-Wei Chang, and Jacob G. Foster. 2022. "Integrating Topic Modeling and Word Embedding to Characterize Violent Deaths." *Proceedings of the National Academy of Sciences* 119(10):e2108801119. doi:10.1073/pnas.2108801119.
- Ashokkumar, Ashwini, and James W. Pennebaker. 2021. "Social Media Conversations Reveal Large Psychological Shifts Caused by COVID-19's Onset across U.S. Cities." *Science Advances* 7(39):eabg7843. doi:10.1126/sciadv.abg7843.
- Ashokkumar, Ashwini, and James W. Pennebaker. 2022. "Tracking Group Identity through Natural Language within Groups." *PNAS Nexus* 1(2):pgac022. doi:10.1093/pnasnexus/pgac022.
- Axelrod, Robert. 1997. *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*. Princeton, NJ: Princeton University Press.
- Baumgartner, Jason, 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* 14:830–39. doi:10.1609/icwsm.v14i1.7347.
- Beauchamp, Zack. 2018. "Incel, the Misogynist Ideology That Inspired the Deadly Toronto Attack, Explained." *Vox*. <https://www.vox.com/world/2018/4/25/17277496/ncel-toronto-attack-alek-minassian>.
- Becker, Howard S. 2008. *Tricks of the Trade: How to Think about Your Research While You're Doing It*. Chicago, IL: University of Chicago Press.
- Berger, Peter L. 1990. *The Sacred Canopy: Elements of a Sociological Theory of Religion*. Washington, DC: National Geographic Books.
- Blumer, Herbert. 1986. *Symbolic Interactionism: Perspective and Method*. Berkeley: University of California Press.
- Choi, Daejin, Jinyoung Han, Taejoong Chung, Yong-Yeol Ahn, Byung-Gon Chun, and Ted Taekyoung Kwon. 2015. "Characterizing Conversation Patterns in Reddit: From the Perspectives of Content Properties and User Participation Behaviors." Pp. 233–43 in *Proceedings of the 2015 ACM on Conference on Online Social Networks, COSN '15*. New York, NY: Association for Computing Machinery.

- DellaPosta, Daniel, Yongren Shi, and Michael Macy. 2015. "Why Do Liberals Drink Lattes?" *American Journal of Sociology* 120(5):1473–511. doi:10.1086/681254.
- Department of Homeland Security. 2022. "Hot Yoga Tallahassee: A Case Study of Misogynistic Extremism." Report. Washington, DC: U.S. Secret Service, Department of Homeland Security.
- DiMaggio, Paul, Clark Bernier, Charles Hecksher, and David Mimno. 2018. "Interaction Ritual Threads: Does IRC Theory Apply Online?" Pp. 81–124 in *Ritual, Emotion, Violence*, edited by E. B. Weininger, A. Lareau, and O. Lizardo. New York, NY: Routledge.
- DiMaggio, Paul, Manish Nag, and David Blei. 2013. "Exploiting Affinities between Topic Modeling and the Sociological Perspective on Culture: Application to Newspaper Coverage of U.S. Government Arts Funding." *Poetics* 41(6):570–606. doi:10.1016/j.poetic.2013.08.004.
- Edelmann, Achim, Tom Wolff, Danielle Montagne, and Christopher A. Bail. 2020. "Computational Social Science and Sociology." *Annual Review of Sociology* 46(1):61–81. doi:10.1146/annurev-soc-121919-054621.
- Farrell, Tracie, Miriam Fernandez, Jakub Novotny, and Harith Alani. 2019. "Exploring Misogyny across the Manosphere in Reddit." Pp. 87–96 in *Proceedings of the 10th ACM Conference on Web Science, WebSci '19*. New York, NY: Association for Computing Machinery.
- Fine, Gary Alan. 1979. "Small Groups and Culture Creation: The Idioculture of Little League Baseball Teams." *American Sociological Review* 44(5):733–45. doi:10.2307/2094525.
- Fine, Gary Alan. 2012. *Tiny Publics: A Theory of Group Action and Culture*. New York, NY: Russell Sage Foundation.
- Geertz, Clifford. 1964. "Ideology as a Cultural System." Pp. 47–76 in *Ideology and Discontent*, edited by D. E. Apter. New York, NY: The Free Press of Glencoe.
- Gibson, David R. 2008. "How the Outside Gets in: Modeling Conversational Permeation." *Annual Review of Sociology* 34(1):359–84. doi:10.1146/annurev.soc.34.040507.134556.
- Gillett, Rosalie, and Nicolas Suzor. 2022. "Incels on Reddit: A Study in Social Norms and Decentralised Moderation." *First Monday* 27(6). doi:10.5210/fm.v27i6.12575.
- Ging, Debbie. 2019. "Alphas, Betas, and Incels: Theorizing the Masculinities of the Manosphere." *Men and Masculinities* 22(4):638–57. doi:10.1177/1097184X17706401.
- Glance, Alyssa M., Tessa L. Dover, and Judith G. Zatzkin. 2021. "Taking the Black Pill: An Empirical Analysis of the 'Incel.'" *Psychology of Men & Masculinities* 22(2):288–97. doi:10.1037/men0000328.
- Goffman, Erving. 1981. *Forms of Talk*. Philadelphia: University of Pennsylvania Press.
- Goffman, Erving. 1986. *Frame Analysis: An Essay on the Organization of Experience*. Boston, MA: Northeastern University Press.
- Goldberg, Amir, and Sarah K. Stein. 2018. "Beyond Social Contagion: Associative Diffusion and the Emergence of Cultural Variation." *American Sociological Review* 83(5):897–932.
- Gonzales, Amy L., Jeffrey T. Hancock, and James W. Pennebaker. 2010. "Language Style Matching as a Predictor of Social Dynamics in Small Groups." *Communication Research* 37(1):3–19. doi:10.1177/0093650209351468.
- Halpin, Michael. 2022. "Weaponized Subordination: How Incels Discredit Themselves to Degrade Women." *Gender & Society* 36(6):813–37. doi:10.1177/08912432221128545.
- Hanson, Kenneth R., C. J. Pascoe, and Ryan Light. 2023. "'It's Getting Difficult to Be a Straight White Man': Bundled Masculinity Grievances on Reddit." *Sex Roles* 88(3):169–86. doi:10.1007/s11199-022-01344-8.
- Hauser, Christine. 2017. "Reddit Bans Nazi Groups and Others in Crackdown on Violent Content." *The New York Times*, October 26.
- Hsu, Kean J., Kalina N. Babeva, Michelle C. Feng, Justin F. Hummer, and Gerald C. Davison. 2014. "Experimentally Induced Distraction Impacts Cognitive but Not Emotional Processes in Think-Aloud Cognitive Assessment." *Frontiers in Psychology* 5:474. doi:10.3389/fpsyg.2014.00474.
- Kaati, Lisa, Amendra Shrestha, and Katie Cohen. 2016. "Linguistic Analysis of Lone Offender Manifestos." Pp. 1–8 in *2016 IEEE International Conference on Cybercrime and Computer Forensic (ICCCF)*. New York, NY: IEEE.
- Kanter, Rosabeth Moss. 1972. "Commitment and Community - Communes and Utopias in Sociological Perspective." <https://www.hup.harvard.edu/catalog.php?isbn=9780674145764>.
- Karell, Daniel, Andrew Linke, Edward Holland, and Edward Hendrickson. 2023. "'Born for a Storm': Hard-Right Social Media and Civil Unrest." *American Sociological Review* 88(2):322–49. doi:10.1177/00031224231156190.
- Lizardo, Omar, and Michael Strand. 2010. "Skills, Toolkits, Contexts and Institutions: Clarifying the Relationship between Different Approaches to Cognition in Cultural Sociology." *Poetics* 38(2):205–28. doi:10.1016/j.poetic.2009.11.003.
- Lofland, John, and Rodney Stark. 1965. "Becoming a World-Saver: A Theory of Conversion to a Deviant Perspective." *American Sociological Review* 30(6):862–75. doi:10.2307/2090965.
- Macy, Michael W., James A. Kitts, Andreas Flache, and Steve Benard. 2003. "Polarization in Dynamic Networks: A Hopfield Model of Emergent Structure." Pp. 162–73 in *Dynamic Social Network Modeling and Analysis*, edited by R. Breiger, K. Carley, and P. Pattison. Washington, DC: National Academies Press.
- Macy, Michael W., Manqing Ma, Daniel R. Tabin, Jianxi Gao, and Boleslaw K. Szymanski. 2021. "Polarization and Tipping Points." *Proceedings of the National Academy of Sciences* 118(50):e2102144118. doi:10.1073/pnas.2102144118.
- Massanari, Adrienne. 2017. "#Gamergate and the Fapping: How Reddit's Algorithm, Governance, and Culture Support Toxic Technocultures." *New Media & Society* 19(3):329–46. doi:10.1177/1461444815608807.
- Nelson, Laura K. 2020. "Computational Grounded Theory: A Methodological Framework." *Sociological Methods & Research* 49(1):3–42. doi:10.1177/0049124117729703.
- O'Donnell, Catharina, and Eran Shor. 2022. "'This Is a Political Movement, Friend': Why 'Incels' Support Violence." *The British Journal of Sociology* 73(2):336–51. doi:10.1111/1468-4446.12923.
- O'Malley, Roberta Liggett, Karen Holt, and Thomas J. Holt. 2022. "An Exploration of the Involuntary Celibate (Incel) Subculture Online." *Journal of Interpersonal Violence* 37(7–8):NP4981–5008. doi:10.1177/0886260520959625.
- Orwell, George. 2021. *Nineteen Eighty-Four*. Penguin Clothbound Classics. London, England: Penguin Classics.

- Pennebaker, James W., Ryan L. Boyd, Kayla Jordan, and Kate Blackburn. 2015. "The Development and Psychometric Properties of LIWC2015." Austin, TX: The University of Texas at Austin.
- Reddit Press. 2023. "Press - Reddit." <https://web.archive.org/web/20230418143443/https://www.redditinc.com/press>.
- Regehr, Kaitlyn. 2022. "In(Cel)Doctrination: How Technologically Facilitated Misogyny Moves Violence off Screens and on to Streets." *New Media & Society* 24(1):138–55. doi:10.1177/1461444820959019.
- Ribeiro, Manoel Horta, Jeremy Blackburn, Barry Bradlyn, Emiliano De Cristofaro, Gianluca Stringhini, Summer Long, Stephanie Greenberg, and Savvas Zannettou. 2021. "The Evolution of the Manosphere across the Web." *Proceedings of the International AAAI Conference on Web and Social Media* 15(1):196–207.
- Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, and David G. Rand. 2014. "Structural Topic Models for Open-Ended Survey Responses." *American Journal of Political Science* 58(4):1064–82. doi:10.1111/ajps.12103.
- Roberts, Margaret E., Dustin Tingley, Brandon M. Stewart, and Edoardo M. Airolidi. 2013. "The Structural Topic Model and Applied Social Science." Paper presented at the Advances in Neural Information Processing Systems Workshop on Topic Models: Computation, Application, and Evaluation, Lake Tahoe, NV. <https://bstewart.scholar.princeton.edu/sites/g/files/toruqf4016/files/bstewart/files/stmnips2013.pdf>.
- Rochford, E. Burke, Jr. 1982. "Recruitment Strategies, Ideology, and Organization in the Hare Krishna Movement." *Social Problems* 29(4):399–410. doi:10.2307/800029.
- Sacks, Harvey. 1992. *Lectures on Conversation*. Oxford, UK: Blackwell.
- Sacks, Harvey, Emanuel A. Schegloff, and Gail Jefferson. 1978. "A Simplest Systematics for the Organization of Turn-Taking for Conversation." Pp. 7–55 in *Studies in the Organization of Conversational Interaction*, edited by J. Schenkein. New York, NY: Academic Press.
- Shi, Qianyi, and Yongren Shi. 2023. "Distinguishing Social Mechanisms of Membership Adoption in Emerging Technology Communities." *Social Science Research* 114: 102917. doi:10.1016/j.ssresearch.2023.102917.
- Smith, Laura G. E., Laura Wakeford, Timothy F. Cribbin, Julie Barnett, and Wai Kai Hou. 2020. "Detecting Psychological Change through Mobilizing Interactions and Changes in Extremist Linguistic Style." *Computers in Human Behavior* 108:106298. doi:10.1016/j.chb.2020.106298.
- Song, Yunya, Qinyun Lin, K. Hazel Kwon, Christine H. Y. Choy, and Ran Xu. 2022. "Contagion of Offensive Speech Online: An Interactional Analysis of Political Swearing." *Computers in Human Behavior* 127:107046. doi:10.1016/j.chb.2021.107046.
- Stark, Rodney, and William Sims Bainbridge. 1980. "Networks of Faith: Interpersonal Bonds and Recruitment to Cults and Sects." *American Journal of Sociology* 85(6):1376–95. doi:10.1086/227169.
- Statista. 2022. "Reddit.Com Desktop Traffic Share 2022." <https://www.statista.com/statistics/325144/reddit-global-active-user-distribution/>.
- Thorburn, Joshua, Anastasia Powell, and Peter Chambers. 2023. "A World Alone: Masculinities, Humiliation and Aggrieved Entitlement on an Incel Forum." *The British Journal of Criminology* 63(1):238–54. doi:10.1093/bjc/azac020.
- Van Valkenburgh, Shawn P. 2021. "Digesting the Red Pill: Masculinity and Neoliberalism in the Manosphere." *Men and Masculinities* 24(1):84–103. doi:10.1177/1097184X18816118.
- Weaver, Matthew, and Steven Morris. 2021. "Plymouth Gunman: A Hate-Filled Misogynist and 'Incel.'" *The Guardian*, August 13.
- Wojcieszak, Magdalena. 2009. "'Carrying Online Participation Offline'—Mobilization by Radical Online Groups and Politically Dissimilar Offline Ties." *Journal of Communication* 59(3):564–86. doi:10.1111/j.1460-2466.2009.01436.x.

Author Biographies

Yongren Shi is a computational sociologist working in the areas of social networks, organizations, social psychology, and culture. He is an assistant professor of sociology at the University of Iowa. The foundation of his research is the sociological study of human behavior and group dynamics. He uses extensively large-scale digital trace data and a wide range of computational methods, including network analysis, computational textual analysis, agent-based computational models, machine learning, online experiments, and sequence analysis. His research has appeared in outlets such as *American Sociological Review*, *American Journal of Sociology*, and *Nature Human Behaviour*, among others. His work has been supported by research grants from the National Science Foundation and covered by popular media outlets such as *Wired*, *The Guardian*, BBC News, *Huffington Post*, and *LA Times*.

Kevin Kiley is a postdoctoral scholar in the Department of Sociology and Anthropology at North Carolina State University. His research explores the social and cognitive sources of opinion and belief stability and change over the life course. His work has been published in the *American Sociological Review* and *Sociological Science*.

Stephanie M. DiPietro is an associate professor in the Department of Sociology and Criminology at the University of Iowa. Her research focuses primarily on life course criminology and patterns of assimilation and behavior among immigrant and refugee groups. Her recent work has been published in *Criminology*, *The British Journal of Criminology*, and *The Annual Review of Criminology*.