

City University of New York (CUNY)

CUNY Academic Works

Student Theses

John Jay College of Criminal Justice

Spring 2-13-2023

Linguistic Predictors of Engagement on Anti-Incel Forums

Daniel Beutler

CUNY John Jay College, Daniel.beutler2122@gmail.com

[How does access to this work benefit you? Let us know!](#)

More information about this work at: https://academicworks.cuny.edu/jj_etds/269

Discover additional works at: <https://academicworks.cuny.edu>

This work is made publicly available by the City University of New York (CUNY).

Contact: AcademicWorks@cuny.edu

Linguistic Predictors of Engagement on Anti-Incel Forums

A Thesis Presented in Partial Fulfillment of the Requirements

For the Master of Arts in Forensic Psychology

John Jay College of Criminal Justice

City University of New York

Daniel Beutler

December 2022

Linguistic Predictors of Engagement on Anti-Incel Forums

Daniel Beutler

This Thesis has been presented to and accepted by the Office of Graduate Studies, John Jay College of Criminal Justice in Partial Fulfillment of the Requirements for the Degree of Master of Science in Master of Arts in Forensic Psychology.

Thesis Committee

Thesis Advisor: Dr. Shuki Cohen

Second Reader: Dr. Charles Stone

External Reader: Dr. Chitra Raghavan

Abstract

This thesis examines the determinants of social media engagement with extremist views in general and misogynistic ones in particular. The study, which is a part of an ongoing, large-scale examination of online hate, examined posts from r/IncelTears, a sub-forum in the long-form social media Reddit dedicated to attacks on the Involuntary Celibate ('Incel') digital communities. Previous research on engagement – and by extension virality – of online posts has broadly implicated linguistic markers of affect, morality, extremism, and social identity. This exploratory study correlated metrics of online engagement such as the number of Upvotes and Comments that each post garnered with its content, as measured by the proportions of a variety of linguistic markers in the body of the post. To ensure reliability in the linguistic analyses, of the 986 posts that were published on the r/IncelTears Subreddit between 05/13/2020 and 06/05/2020 only 438 posts that exceeded 100 words were included in our final analysis.

Our findings largely concur with previous research concerning the key role that (negative) affectivity and (indignant) morality, as well as ingroup/outgroup 'tribalism' markers play in driving the users' public reactions to online posts. Additionally, the current study revealed linguistic markers that differentiate – at times substantially – between low-level engagement (e.g. Upvotes) and high-level engagement (e.g. Comments). Novel, or minimally-studied linguistic drivers of engagement were also identified, including the use of pseudo-scientific language (e.g. the use of words over 6 letters, comparison words, and references to biology and females).

The current study, which is a part of a larger endeavor to understand the online hate ecosystem, is based on a snapshot of less than a month-worth of posts in one of the most prolific cross-platform online hate causes. Further research is needed to further elucidate the linguistic

‘hot buttons’ that drive engagement with hate speech, to prioritize their subjective salience and import, and ultimately to construct effective online counter-messaging or de-escalation campaigns to mitigate this ubiquitous problem.

Keywords: Incels, Social Media, Psycholinguistics, Engagement, Reddit

Table of Contents

Abstract.....	3
Table of Contents.....	5
Introduction.....	6
Incels.....	6
Emotional Language and Social media.....	8
Reddit.....	13
Current Study.....	14
Background.....	14
Methods.....	15
Research Design.....	15
Procedure.....	16
Results.....	17
Discussion.....	20
Limitations and future research.....	22
Conclusion.....	24
Bibliography.....	26

Introduction

Incels

Much has been said in recent years about the group of people who refer to themselves as involuntarily celibate or incels. Incels, while largely non-violent (Speckhard, Ellenberg, Morton & Ash, 2021), are infamous for homicide events. Members of the incel community have committed several mass homicide events, including the Isla Vista Shooting (Witt, 2020) and the Toronto bus attack ("Alek Minassian Toronto van attack suspect praised 'incel' killer", 2021).

Incels predominantly meet on online forums to discuss grievances regarding sexual dysfunction and social dysfunction. These grievances tend to be in regard to their inability to find sexual partners. Ideologically at the surface, incels believe in central themes of misogyny, and male oppression and seek to legitimize masculinity (O'Malley, Holt & Holt, 2020). Inceldom by all accounts is a growing phenomenon with web traffic to popular incel forums growing six times during one six-month period in 2021 (Ball, 2022). The most popular incel forum garners around 2.6 million visits a month from around 11,000 active users with an average of around 170,000 posts a month (Center for Countering Digital Hate, 2022).

Amongst Incels there is no consensus of what constitutes membership, with some incels requiring adherence to a complex set of ideological prescriptions known as "the blackpill" for one to be considered an incel.(Speckhard, Ellenberg, Morton & Ash, 2021) whereas others believe that inceldom is defined simply by an inability to garner sexual companionship (Incel Wiki, 2022).

Incels who follow the blackpill believe that in order to be a member of the incel community one must be a part of a genetic underclass of men who have low value in the

perceived sexual marketplace. They believe that one's value in society is entirely determined by a genetic lottery that separates the unattractive men, Incels, average-looking men, “Normies”, and attractive men, “Chads”. This hierarchical difference is made clear by the lack of attention Incels perceive to receive from women who they refer to as “Stacy” (Incel Wiki, *Blackpill* 2022). The grievance they feel in regards to the genetic lottery goes further than just sexual relationships, with blackpillers believing that they are mistreated by employers and society as a whole due to their genetic inferiority. Amongst blackpilled Incels, it is a common belief that the genetic hierarchy is enforced by Jewish leaders in media and government, who use cultural movements such as #metoo and feminism broadly in order to control and subjugate incels. This is a broad contrast to less extreme incels who adopt the “redpill”, who don't believe that they are genetically unable to succeed while still broadly believing in the hierarchy between Chad, Stacy, and Normies (Hoffman, Ware, & Shapiro. 2020). Inside the incel community, the more extreme blackpilled users believe that redpillers are in denial or are not actual incels.

Due to the somewhat contentious definition of inceldom, some researchers have defined the minimum requirements of inceldom as “ being male, and not having had a sexual partner for a long time” (Jaki et al., 2019). This is perhaps the most permissive definition of what makes an incel, painting with a fairly broad brush.

Inceldom does not exist in a vacuum, In fact, inceldom is one section of a broadly misogynistic group of internet users that have been dubbed by researchers as the “manosphere”. Groups within the manosphere include men's rights activists, pick-up artists, and have a degree of overlap with more overtly political entities such as the proud boys (Hoffman, Ware, & Shapiro. 2020). Like other communities in the manosphere, incel forums are seeing continuous and rapid growth (Farrell, Fernandez, Novotny & Alani, 2019). What is notable about incels is

that they seem to use higher levels of hate speech even when compared to other expressly misogynist online forums. (Ribeiro, et al, 2021). In order to gain a better handle on the way inceldom is growing, this analysis intends to look at the emotional language that is used on these forums and the effect it has on engagement.

Emotional Language and Social Media

With the ubiquitous nature of social media, attention must be paid to the manner in which platforms impact their users. Given that platforms such as Facebook boast numbers of nearly 3 billion users (Iqbal, 2022), it's hard to overstate the impact they have on society. It has been shown that social media websites like Twitter have a material impact on the results of elections (Fujiwara, Müller, & Schwarz, 2021) and internal documents at Facebook reveal that Instagram has a negative effect on the body image of young women (Wells et al., 2021). This is all to say the perhaps obvious, that the media that we consume has an effect on us. Therefore it would be prudent to examine the underlying mechanisms that drive social media and the ideas that are spread with it.

Looking at advertising research we can find that both positive and negative language have been shown to drive engagement. For the sake of this analysis we will be defining engagement as anytime a user takes a deliberate action to interact with a piece of online content. The most common forms of digital engagement are comments, likes and shares. With it being believed that positive affective language increases sentimentality and therefore engagement; conversely negative language is believed to drive polarity which increases engagement (Arapakis, Lalmas, Cambazoglu, Marcos & Jose, 2014).

Online content has a tendency to self-propagate via both peer-to-peer sharing and large broadcasters (e.g. News papers, Political pundits, Social commentators) highlighting specific posts or trends. When content diffuses between users they tend to do so sparingly. For example, when sharing or resharing content via social media 99% of posts are shared either directly from the original website or a person who was resharing it from the original website (Goel et al., 2012). This tendency for individual posts to not cascade down too far from the source is true for both posts shared by broadcasters and individuals (Goel et al., 2016). This could mean that it is not the broadcaster or individual sharing the post but rather some quality of the post itself that drives engagement.

Research on the subject reveals both the emotionality of the post and the specifics of the wording of the post are key components in driving engagement. In regards to emotionality, posts that are high in arousal language such as anger tend to be more viral, as well as posts that invoke feelings of dominance such as happiness or inspiration (Guerini & Staiano, 2015). The specific wording of posts also seems to matter a great deal, with ingroup language and references to people other than yourself and the reader being linked with increased virality. Words like, “we”, “us” and “I” are indicative of ingroup language (Tan et al., 2014).

Emotion seems to propagate much in the same way that content does on social media networks. This was shown clearly by the Facebook emotional contagion study in which nearly 700,000 users were presented with different curated emotional content on their Facebook feed without their knowledge. It was found that the subsequent posts by the user were more likely to present the same emotion the user was shown. (Kramer, Guillory and Hancock, 2014). This study may provide a hint towards the mechanism by which content propagates. Sharing content

could be a way in which one signals their heightened emotionality from the very content that they share.

The effects of content propagation can be seen very clearly by research into the so-called “alt-right pipeline” which has revealed a kind of snowball effect in which users are recommended more extreme content from relatively benign viral content (Ribeiro, et al, 2020). We have reason to believe that the manosphere operates similarly due to its extremist belief system and a large degree of overlap with the “alt-right” (Mamié, Ribeiro, & West, 2021). Extremist language is also effective at driving follower counts; the most politically extreme members of the House of Representatives have significantly more followers than their less extreme colleagues (Lam & Hughes, 2022). Of the social media communication styles used by politicians, an emotional style has been found to be the most effective in driving social media engagement (Keller & Kleinen-von Königslöw, 2018). Engaging with political leaders has a major impact on not only their policy decisions but also on the level of the political activity of those who engage. People who engage with political posts are much more politically knowledgeable and active than people who simply consume news content (Dimitrova, Shehata, Strömbäck, & Nord, 2011). Taking all of this together we see that more extreme political content is effective at increasing the reach of political ideology and the impact that ideology has on those who follow it.

More extreme political posts driving activity may have some terrible repercussions when consideration is given to the fact that social media posts that are in reference to outgroups are more likely to draw engagement, with one study finding that including these references will increase the likelihood of a tweet being shared by 68%. That same study found that sharing outgroup-referenced posts led to an increase in negative comments. (Rathje, Van Bavel & van

der Linden, 2021). In a vacuum, this points towards extremist content being directed towards others as an effective way to disseminate a message but this can also be seen as a way in which violence is directed. The link between political rhetoric and extreme action such as violence can be shown in a number of ways, such as terrorist organizations like ISIS's adoption of social media as a recruitment tool. Another example is the rise of anti-Asian sentiment online during the COVID-19 pandemic (Nguyen et al., 2020) which coincided with an increase in violence on the Asian Americans in the United States (Gill et al, 2019) .It is worth noting that this rise of anti-Asian American sentiment did also coincide with a large counter-hate movement (Ziems, He, Soni & Kumar, 2020) Research on online hate groups found a positive correlation between the size of the group and the amount of negative emotional content the group produced. For example, the subreddit r/altright, used hate speech at a rate much higher than similar-sized political forums on the same platform while seeing a large amount of growth (Grover, Mark, 2019). The subreddit garnered a large amount of attention in both the media and by researchers.

One type of language that seems to have a notable impact on the rate at which a post is shared or diffuses through a community is “Moral Emotional” language. Each inclusion of a moral emotional word such as “hate” increases the rate at which a post is retweeted by 20% on average. This is in contrast to strictly moral words like “duty” which seem to have no significant effect on the rate in which posts are shared on Twitter. Words like “fear” that are distinctly emotional do increase the diffusion rate of a post but not nearly as much as moral emotional language. Another finding of this same study is that Moral Emotional language tends to diffuse more easily through political groups when compared to posts that are viewed by members outside of the poster's intended audience (Brady et al., 2017). A major contributing factor to the diffusion power of moral emotional language seems to be the ability to capture the attention of

the reader. It has been found that moral and emotional language both contribute are more likely to capture the attention of a reader than neutral language, furthermore, it has been found that attention is related to resharing. Interestingly the relationship between moral and emotional language and diffusion seems to be independent of the impact of arousal on diffusion (Brady et al., 2019).

This tendency for moral or emotional language to impact the level of diffusion an online post has can have real-world implications for the authorship of effective political messaging. A study looking at messaging in regard to the COVID-19 pandemic found that deontological language which it defined as messaging in regards to the “right thing to do” was much more effective than non-moral language at getting research subjects to want to reshare content. The study also found that deontological language is much more impactful in regard to content diffusion when the message is coming from a political authority figure (Everett et al., 2020). The ability of moral language to impact sharing behavior and therefore content diffusion is not merely an English phenomenon. A study on Chilean social media sharing found that morally framed posts were by far more likely to be shared when compared to posts framed in more generic ways. That same study also found that posts that were framed around conflict were less likely to be shared on both Facebook and Twitter. Interviews with study participants revealed that they didn’t want to share news that they found divisive and preferred to share moral posts that were seen as more unifying (Valenzuela et al., 2017).

As knowledge of inceldom becomes more mainstream so do communities whose express purpose is to make fun of, catalog, and talk about incels. The members of these communities scour the internet for the incel posts that they believe to be the most interesting. These communities are valuable in many ways, they are useful for researchers because they curate and

gather the most hateful posts on various incel and manosphere communities and they are useful for the public as a way to effectively disseminate counter-messaging. It has been shown that you can mitigate the persuasive effect of extremist content by inoculating users with counter-messaging before they view the content. Showing someone counter-messaging before extremist content lowers the perceived credibility of the extremist group, increases psychological resilience, and reduces the intent to support the extremist group. These effects are regardless of the source of the inoculating message (Braddock, 2020). Inoculation is effectively built into forums such as Reddit's r/inceltear, on inceltear posts are hosted via third-party image services with only the title being visible on the Reddit client. This necessitates that users see the title before viewing the content, with the titles often being refutations or mockeries of the extreme content of the post. What this means is that users are shown effective counter-messaging before ever experiencing any extremist content.

In looking at the posts that these communities gather from various incel forums, we are able to see some of the most viral and controversial posts inceldom has to offer. We hope that by looking at these posts we can get a snapshot of the emotional language used by incels. This will provide us with a better understanding as to why some posts go more viral than others, which is valuable because it gives us insights into the radicalization process.

Reddit

Self-styled as the “front page of the internet”, Reddit.com is a widely popular social media website founded in 2005 with over 440 million active monthly users (Curry, 2022). Reddit users or Redditors spend their time on smaller forums that are known as subreddits. Subreddits are user-created and user moderated. This has led to the creation and curation of various different

communities all centered around user interest. Broadly users on Reddit interact with each other in four different ways: voting, comments, awards, and direct messages. Votes act as endorsements that content is quality and deserves to be pushed to more users' Reddit feeds, users can upvote posts they believe to be quality and downvote posts that they feel are of poor quality. Users can also pay Reddit to give awards to posts that they feel strongly about, this increases the number of users who see a post. Users can also comment on posts, these comments are sorted by the aggregate of upvotes and downvotes like other content on the website. One can also directly message another user but as they are private, these will not be examined in the study. Reddit has a history of political activism with subreddits for supporters of then-candidate Donald Trump garnering nearly 800,000 subscribers. The /r/TheDonald subreddit has since been banned but other subreddits like /r/sandersforpresident and /r/conservative still retain large footholds on the platform.

Current Study

Background

The purpose of this study was to look at how the language of incels drives engagement on anti-incel forums. To do this we utilized the Reddit forum, r/inceltear. This forum allows non-incel users to post various incel and misogynic content from around the internet. At the time of this study, r/inceltear has 112,360 users. Content on r/inceltear is made up nearly completely of screen grabs of text-based posts from incel forums both on Reddit and on their own dedicated websites. Users of r/inceltear scour the internet looking for the most incendiary incel content that they can find. Broadly we believe that the level of engagement on r/inceltear will be similar to the level of engagement that incels would present as the interests of their communities align.

Posts that are successful on r/incel are likely also successful on the incel forums that the content is drawn from. This is inline with research that suggests that anti-extremist groups often have similar interests and will even mimic the extremist groups that they are in opposition too. The delineation between actively against and being curious about extremist ideology is understood to be a fuzzy barrier. This can lead to anti-incels adopting a lot of the tactics that incels use. (Ingram, 2016). This mimicry is perhaps even tighter due to the short window of relevance that social media affords.

Data for this study was taken in the form of screen captures of the “new” section of r/incel over the course of a one month time period. This has resulted in a complete catalog of user-submitted posts. Using r/incel posts also allows us to take user engagement data directly from Reddit, something that would be impossible on other online forums without publicly facing engagement metrics like the incel forums that the posts are drawn from. Looking at posts from incel forums provides us with a unique perspective on extremist content, due to a number of factors. The first is their aforementioned high use of hate speech combined. Perhaps the most unique part of inceldom as an extremist political movement is their reliance on publicly facing text-based communication, this is very unlike hate groups such as the Ku Klux Klan or ISIS which utilize in-person meetings, private encrypted chat channels and video in order to disseminate their message. This study stands apart from current research in its scope, with most research on social media studies being focused on more macro level trends. Looking at a highly politicized group allows us to see the key factors to engagement even on the fringes of the web.

Methods

Research Design

The present study is an exploratory, correlative study. Using the various scales from LIWC-15 and LIWC-22, we explored the relationship between language and social media engagement. We opted to use both LIWC-15 and LIWC-22 for a number of reasons. First and foremost LIWC-15 and LIWC-22 look at affect in different ways. LIWC-22 creates a split between emotional language and sentimental language. LIWC-22 restricts emotional words as “true emotion labels, as well as words that strongly imply emotions” and sentiment words as “reflecting sentiment rather than emotion per se”, words like “birthday” and “funeral” would be considered tone words indicating sentiment (Boyd, et al 2022). LIWC-15 does not make a distinction between the component's that makeup affect, which allows for a much more general measure of affect (Pennebaker, et al 2015). LIWC-15 is also a much more tested and established measure so we thought it appropriate to rely on it for the majority of our analysis. On the other hand, LIWC-22 aligns much more closely with current research in regards to the difference between affective tone and emotion. For those reason, we used LIWC-15 for our more general measures and LIWC-22 to measure specifically tone.

Procedure

Over the course of the months of June and May 2020, every post and its corresponding engagement metrics from reddit's /r/incel subreddit were cataloged. These posts were largely in the form of screen captures from various involuntary celibate forums or adjacent ideological forums. These posts were sorted between higher and lower image quality based on file size. Lower-quality images were manually transcribed, by student research assistants trained to be familiar with the involuntary celibate community and its linguistics. All research assistants that provided transcription services were trained by Dr. Shuki Cohen and are participants in his Far

Right Extremism Lab. The posts that comprise this study are one small part of a larger data set of incel posts being compiled by Dr. Cohen's lab. The posts that are used are from a randomly selected timeframe of one month out of the 45 months of data that the lab currently has.

Higher image-quality posts were fed through the Optical character recognition (OCR) software built into the Adobe Acrobat PDF reader software. The transcribed posts from the OCR were then looked over to ensure accuracy. Finally, the text data are run through LIWC-15 and LIWC-22 to extract the emotional coding of the original posts.

Results

The current study looked at 1053 posts from reddit's r/incel subreddit. 616 of those posts were excluded due to a number of factors including: being non-textual, having a word count under 200 words, or a lack of engagement data. This left us with a database of 437 posts with an average length of 233 words per post for an overall database of 102,312 words. On average a post received 52.21 upvotes with a standard deviation of 97.36, Posts were commented on at a rate of 19.33 on average with a standard deviation of 22.410.

Overall engagement on incel forums seems to be correlated with a number of metrics. First and foremost, upvotes and comments are highly correlated ($r=0.842$, $p<0.0001$). General references to affect positively drives engagement (upvotes: $r=0.103$, $p<0.05$; comments: $r=0.117$, $p<0.05$), this seems to be in part driven by positive sentimentality and emotion which are both significantly correlated with comments (emotion: $r=0.102$, $p<0.05$; sentimentality: $r=0.104$, $p<0.05$) but not necessarily upvotes (emotion: $r=0.072$, $p>0.05$, sentimentality: $r=0.089$, $p>0.05$).

Both engagement indicators are positively correlated with: proportion of pronouns (upvotes: $r=0.100$, $p<0.05$; comments: $r=0.107$, $p<0.05$), especially personal pronouns (upvotes:

$r=0.111$, $p<0.05$; comments: $r=0.132$, $p<0.01$), first person pronouns (e.g. "I"; upvotes: $r=0.124$, $p<0.01$; comments: $r=0.123$, $p<0.01$). Engagement is also positively correlated with the use of third person singular pronouns (upvotes: $r=0.120$, $p<0.05$; comments: $r=0.110$, $p<0.05$). Third person pronouns being related to engagement is likely due to references to female third person pronouns (e.g. "She/Her," upvotes: $r=0.141$, $p<0.01$; comments: $r=0.098$, $p<0.05$) rather than references to male third person pronouns (e.g. "he/him" upvotes: $r=-0.049$, $p>0.3$; comments: $r=-0.09$, $r=-0.027$, $p>0.5$). Third person plural pronouns e.g. "They", were associated with less upvotes ($r=-0.096$, $p<0.05$) but had no effect on comments ($r=-0.070$, $p>0.1$).

The prevalence of comparison markers negatively impacts engagement (upvotes: $r=-0.105$, $p<0.05$; comments: $r=-0.130$, $p<0.01$). Engagement was positively correlated with words associated with 'home' (e.g. kitchen, landlord, etc) (upvotes: $r=0.125$, $p<0.01$; comments: $r=0.181$, $p<0.01$). Words over 6 letters were associated with less engagement in the form of comments ($r=-0.110$, $P<0.05$) but had no effect on upvotes ($r=-0.074$, $p>0.1$). references to body parts and function are associated with more comments ($r=0.126$, $p<0.05$) but only marginally with more upvotes ($r=0.091$, $p<0.05$), while general reference to biology are associated with an increased engagement through both more upvotes ($r=0.112$, $p<0.05$) and comments ($r=0.126$, $p<0.01$). Quantitative markers (incl. numbers) and references to sex are associated with more upvotes ($r=0.102$, $p<0.05$) but not significantly with more comments ($r=0.062$, $p>0.05$).

Posts with polite words are linked to more upvotes ($r=0.111$, $p<0.01$), more comments ($r=0.12$, $p<0.01$), likewise prosocial posts were also linked with more upvotes ($r=0.104$, $p<0.05$) and more comments ($r=0.102$, $p<0.05$).

Figure 1. LIWC Upvotes Significant correlations

Category	Example	Pearson's R
Comments	Number of comments	0.842
Female references	girl, her, mom	0.141
Home	kitchen, landlord	0.125
1st person singular	I, me, mine	0.122
3rd person singular pronouns	she, her, him	0.120
Biological processes	eat, blood, pain	0.112
Personal pronouns	I, them, her	0.111
Politeness*	thank, please, thanks, good morning	0.111
Physical*	medic, food, patients, eye	0.110
Prosocial Behavior*	care, help, thank, please	0.105
Affective processes	happy, cried	0.103
Sexual	horny, love, incest	0.102
Articles	a, an, the	-0.095
3rd person plural	they, their, they'd	-0.096
Comparisons	greater, best, after	-0.105
Quantifiers	few, many, much	-0.111

Figure 2. LIWC Comments Significant Correlations

Category	Example	Pearson's R
Home	kitchen, landlord	0.181

Personal pronouns	I, them, her	0.132
Politeness*	thank, please, thanks, good morning	0.126
Body	cheek, hands, spit	0.126
1st person singular	I, me, mine	0.123
Physical*	medic, food, patients, eye	0.120
Affective processes	happy, cried	0.117
Biological processes	eat, blood, pain	0.110
3rd person singular pronouns	she, her, him	0.110
Total pronouns	I, you, that, it	0.107
Positive tone*	good, well, new, love	0.104
Positive emotion*	good, love, happy, hope	0.103
Prosocial Behavior*	care, help, thank, please	0.102
Female references	girl, her, mom	0.098
Words > 6 letters	-	-0.110
Comparisons	greater, best, after	-0.130

*LIWC 22 dimension

Discussion

Looking at 102,312 words from Reddit's r/incel forum we have found a number of interesting relationships between content and engagement. On the surface level, we can be assured that the characteristic of a post relates to engagement with the post. Perhaps the most salient point is that emotionality seems to be a key driver in engagement. Looking a bit deeper we find that positive sentimentality and emotion drive comments but not upvotes, whereas negative sentimentality alone does not seem to be a significant driver for either upvotes or comments. It seems that affect in general is a good driver of engagement but slicing into positive

or negative creates a bit too much noise in the data. It's possible that using a larger data set one would be able to find more of a relationship between positive or negative affect and engagement.

The difference between engaging in the form of comments and in the form of upvotes appears a few times in our findings. Fundamentally comments and upvotes are very different; upvotes are simply an endorsement of the content. Alternatively, comments are much more diverse; they can be used as a way to refute or endorse the content or they could be entirely divergent thoughts ranging from humor, to community support, to completely unrelated discourse. Another major difference between comments and upvotes lies in the effort required to engage, with comments requiring a fair bit more effort than upvotes. This effort is compounded by the emotional effort of opening oneself up to criticism by other comments, this is perhaps especially true in a community like r/incel that deals with sensitive subject matter. When we consider the upvote comment divide in positive sentiment under this lens, we can see that perhaps positive sentiment posts are less challenging posts to comment on while perhaps not being attention-grabbing enough to warrant upvotes.

We see an interesting effect of large words on posts, where it depresses the number of comments but not the number of upvotes. This may speak to an approach-avoidance conflict, broadly speaking an approach-avoidance conflict happens when there are both positive and negative results of an action that make taking an action simultaneously appealing and unappealing (Lewin, 1936). Commenting on a post with complex syntax requires much more out of the commenter and comes with a higher degree of risk. Simply put, refuting a post like that takes more energy and opens the commenter up to a higher degree of criticism. We can use this line of reasoning to explain why we don't see any effect on upvotes, the complexity of the post for most users is not enough to change their low-cost upvoting behavior.

This analytical modality could also be used to explain why users are more likely to upvote and not comment on posts that are in reference to sex. On the positive side, a commenter has very little to gain by affirming the widely held belief by the community. Looking at the negative, by commenting users are engaging with extremely misogynistic content that has a good chance to be very graphic. Commenters also open themselves up to criticism if they refute the post in a way that is different than the community norms. Overall commenting on posts about sex maybe not worth the effort for most users but the much simpler upvote may be.

We see a fair bit of statistical evidence for the power of pronouns to drive engagement. In particular, we see feminine pronouns as the most impactful; this lines up very closely with current research that holds that outgroup references seem to be effective at driving engagement. This is of particular note when you consider who engages with the posts on r/incel; non-incels. Personal pronouns driving engagement can be interpreted in another way when you consider that quantification language, and comparison language also drives engagement. This could be evidence that tribal us vs them style language is perhaps more indicative of a more engaging post rather than simple references to outgroups.

Being as inceldom is an ideology that revolves around sex, it should come as no surprise that there is a relationship between sex and engagement at least as far as upvotes are concerned. What is a bit more interesting is that references to biology seem to have a greater effect. This seems to be impacted in a large part by the prevalence of the types of language in incel discourse, with incels referring to sexual acts often times in more biological terms rather than the active words that LIWC would be looking for in regards to sex. This can be shown pretty clearly by the post that scored the highest on the biology score in which an incel forum poster said “your hormones will have a massive lust for physical contact with the opposite sex” in regards to

teenage development. That post scored a relatively middling score in the sexual battery despite it being entirely about sex. What this suggests is the the biological battery on LIWC may be a better way to evaluate how sexual references drive engagement in this context.

Limitations and future research

It is impossible to interpret this data without speaking about the large degree of variance between posts in terms of engagement. The top posts are being engaged with hundreds and in some cases thousands of times, but most posts garner rather meager engagement metrics. This is both a limitation and entirely expected when dealing with a social media platform like reddit. By design social media websites like reddit promote content that is drawing more attention. This varies from platform to platform with reddit primarily using upvotes as the metric that drives posts in front of users' eyes. With that said, users are able to boost posts by the purchase of reddit awards, awards are intended to be given to other users and not be used to promote one's own content. To that end, users are not able to purchase awards on their own behalf, but this is easily overcome with a second account. This all is to say that engagement often snowballs on advertising-supported attention-based social media websites like reddit so a large degree of variance in engagement metrics is to be expected.

Further limitations include what we are using for the upvote engagement, the publicly facing score that we have been using to measure upvotes is actually an aggregate score of all upvotes and downvotes. This if anything means that our engagement metrics are understated but it would be impossible to say to what degree without private information only available to reddit

administrators or the subreddit's moderation team. Unfortunately, there's very little that can be done to mitigate this limitation without directly working with website stakeholders.

There is a litany of research that may be done to expand upon the ideas explored in this thesis. First and foremost explaining this style analysis to other communities in order to examine if emotional language drives engagement across the political spectrum. It's also highly worthwhile to further explore engagement's relationship to political action, perhaps the easy way to measure this would be voting behavior. Another avenue of research that could be explored is the impact that content moderation has on engagement; forums unilaterally have both implicit and explicit rules about what can be said and this surely has an effect on the engagement on those forums.

Of course, another fruitful avenue of research would be to run the same kind of analysis using a larger data set with other linguistic batteries such as the moral emotional language battery developed by the Van Bavel lab (Brady, 2017) or a model that is specific to the social media platform that is being researched. For the purpose of this study we felt that it was non-ideal to run our data through these batteries due to the lack of incel specific dictionary and the relatively small sample size. The one month snapshot of inceltar that this was taken from the large data set also does not guarantee the stability that we would like to see to make these types of analysis.

Conclusion

The current study adds to the existing knowledge in social media studies by exploring the way in which language impacts engagement. Through a thorough examination of Reddit's /r/inceltar community we have found key factors which seem to drive engagement. Emotional language in particular positive affective language has a markable effect on the traffic that a post receives. This study's findings are by in large in agreement with the current research findings on

similar topics. This has implications for social media studies, counter extremism and communications in general. The findings of this study illustrate ways in which content moderation can be directed and counter messaging developed in order to decrease the prevalence of online hate content.

Bibliography

- Alek Minassian Toronto van attack suspect praised 'incel' killer. (2021). Retrieved 28 October 2021, from <https://www.bbc.com/news/world-us-canada-43883052>
- Arapakis, I., Lalmas, M., Cambazoglu, B., Marcos, M., & Jose, J. (2014). User engagement in online News: Under the scope of sentiment, interest, affect, and gaze. *Journal Of The Association For Information Science And Technology*, 65(10), 1988-2005. doi: 10.1002/asi.23096
- Ball, T. (2022, January 3). *Massive rise in use of incel sites that call for women to be raped*. News | The Times. Retrieved November 12, 2022, from <https://www.thetimes.co.uk/article/massive-rise-in-use-of-incel-sites-that-call-for-women-to-be-raped-hddbq5mgc>
- Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). The development and psychometric properties of LIWC-22. *Austin, TX: University of Texas at Austin*.
- Braddock, K. (2020). Vaccinating against the enemy: Attitudinal inoculation, radicalization, and counter-radicalization. *Weaponized Words*, 110–137. <https://doi.org/10.1017/9781108584517.007>
- Brady, W., Wills, J., Jost, J., Tucker, J. and Van Bavel, J., 2017. Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*, 114(28), pp.7313-7318.

Brady, W. J., Gantman, A. P., & Van Bavel, J. J. (2019). Attentional capture helps explain why moral and emotional content go viral. <https://doi.org/10.31234/osf.io/zgd29>

Center for Countering Digital Hate. (2022, September 23). *THE INCELOSPHERE Exposing pathways into incel communities and the harms they pose to women and children.*

Retrieved from

<https://counterhate.com/wp-content/uploads/2022/09/CCDH-The-Incelosphere..pdf>

Curry, D. (2022, September 6). *Reddit revenue and Usage Statistics (2022)*. Business of Apps.

Retrieved November 24, 2022, from

<https://www.businessofapps.com/data/reddit-statistics/>

Dimitrova, D., Shehata, A., Strömbäck, J. and Nord, L., 2011. The Effects of Digital Media on Political Knowledge and Participation in Election Campaigns. *Communication Research*, 41(1), pp.95-118.

Everett, J. A., Colombatto, C., Chituc, V., Brady, W. J., & Crockett, M. (2020). The effectiveness of moral messages on public health behavioral intentions during the covid-19 pandemic.

<https://doi.org/10.31234/osf.io/9yqs8>

Farrell, T., Fernandez, M., Novotny, J., & Alani, H. (2019, June). Exploring misogyny across the manosphere in reddit. In *Proceedings of the 10th ACM Conference on Web Science* (pp.

87-96).

Fujiwara, T., Müller, K., & Schwarz, C. (2021). *The effect of social media on elections: Evidence from the United States* (No. w28849). National Bureau of Economic Research.

Gill, P., Braddock, K., Zolghadriha, S., Rottweiler, B., & Cushenbery, L. (2019). *Terrorist Propaganda after IS: Learning, Emulation, and Imitation*. Oxford University Press.

Goel, S., Watts, D. J., & Goldstein, D. G. (2012). The structure of Online Diffusion Networks.

Proceedings of the 13th ACM Conference on Electronic Commerce - EC '12.

<https://doi.org/10.1145/2229012.2229058>

Goel, S., Anderson, A., Hofman, J., & Watts, D. J. (2016). The structural virality of online

diffusion. *Management Science*, 62(1), 180–196. <https://doi.org/10.1287/mnsc.2015.2158>

Guerini, M., & Staiano, J. (2015). Deep Feelings: A Massive Cross-Lingual Study on the

Relation between Emotions and Virality. *Proceedings of the 24th International*

Conference on World Wide Web. <https://doi.org/10.1145/2740908.2743058>

Grover, T., & Mark, G. (2019). Detecting Potential Warning Behaviors of Ideological

Radicalization in an Alt-Right Subreddit. *Proceedings of the International AAAI*

Conference on Web and Social Media, 13(01), 193-204. Retrieved from

<https://ojs.aaai.org/index.php/ICWSM/article/view/3221>

Hoffman, B., Ware, J., & Shapiro, E. (2020). Assessing the threat of incel violence. *Studies in*

Conflict & Terrorism, 43(7), 565-587.

Incel Wiki (2022) *Incel*, *Incel Wiki*. Incel Wiki. Available at: <https://incels.wiki/w/Incel>

(Accessed: November 9, 2022).

Incel Wiki. (2022, November 18). *Blackpill*. Incel Wiki. Retrieved from

<https://incels.wiki/w/Blackpill>

Ingram, H. (2016). A “linkage-based” approach to combating militant Islamist propaganda: A

two-tiered framework for practitioners. *Terrorism and Counter-Terrorism Studies*.

<https://doi.org/10.19165/2016.2.06>

Iqbal, M. (2022, August 31). *Facebook revenue and Usage Statistics (2022)*. Business of Apps.

Retrieved November 13, 2022, from

<https://www.businessofapps.com/data/facebook-statistics/>

Lam, O. and Hughes, A., 2022. Highly ideological members of Congress have more Facebook followers than moderates do. [online] Pew Research Center. Available at:

<<https://www.pewresearch.org/fact-tank/2017/08/21/highly-ideological-members-of-congress-have-more-facebook-followers-than-moderates-do/>> [Accessed 28 April 2022].

Lewin, K. (1936). A dynamic theory of personality: Selected papers. *The Journal of Nervous and Mental Disease*, 84(5), 612-613.

Jaki, S., De Smedt, T., Gwózdź, M., Panchal, R., Rossa, A., & De Pauw, G. (2019). Online hatred of women in the Incels.me forum. *Journal Of Language Aggression And Conflict*, 7(2), 240-268. doi: 10.1075/jlac.00026.jak

Keller, T. and Kleinen-von Königslöw, K., 2018. Pseudo-discursive, mobilizing, emotional, and entertaining: identifying four successful communication styles of political actors on social media during the 2015 Swiss national elections. *Journal of Information Technology & Politics*, 15(4), pp.358-377.

Kramer, A., Guillory, J. and Hancock, J., 2014. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24), pp.8788-8790.

Mamié, R., Horta Ribeiro, M., & West, R. (2021, June). Are Anti-Feminist Communities Gateways to the Far Right? Evidence from Reddit and YouTube. In *13th ACM Web Science Conference 2021* (pp. 139-147).

Nguyen, T. T., Criss, S., Dwivedi, P., Huang, D., Keralis, J., Hsu, E., Phan, L., Nguyen, L. H.,

Yardi, I., Glymour, M. M., Allen, A. M., Chae, D. H., Gee, G. C., & Nguyen, Q. C.

(2020). Exploring U.S. shifts in Anti-Asian sentiment with the emergence of covid-19.

International Journal of Environmental Research and Public Health, 17(19), 7032.

<https://doi.org/10.3390/ijerph17197032>

O'Malley, R., Holt, K., & Holt, T. (2020). An Exploration of the Involuntary Celibate (Incel)

Subculture Online. *Journal Of Interpersonal Violence*, 088626052095962. doi:

10.1177/0886260520959625

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and*

psychometric properties of LIWC2015.

Rathje, S., Van Bavel, J. and van der Linden, S., 2021. Out-group animosity drives engagement

on social media. *Proceedings of the National Academy of Sciences*, 118(26).

Ribeiro, M. H., Blackburn, J., Bradlyn, B., De Cristofaro, E., Stringhini, G., Long, S., ... &

Zannettou, S. (2020). From pick-up artists to incels: a data-driven sketch of the

manosphere.

Ribeiro, M. H., Ottoni, R., West, R., Almeida, V. A., & Meira Jr, W. (2020, January). Auditing

radicalization pathways on YouTube. In *Proceedings of the 2020 conference on fairness,*

accountability, and transparency (pp. 131-141).

Tan, C., Lee, L., & Pang, B. (2014). The effect of wording on message propagation: Topic- and

author-controlled natural experiments on Twitter. *Proceedings of the 52nd Annual*

Meeting of the Association for Computational Linguistics (Volume 1: Long Papers).

<https://doi.org/10.3115/v1/p14-1017>

Speckhard, A., Ellenberg, M., Morton, J., & Ash, A. (2021). Involuntary Celibates' Experiences of and Grievance over Sexual Exclusion and the Potential Threat of Violence Among

Those Active in an Online Incel Forum. *Journal Of Strategic Security, 14*(2), 89-121.

doi: 10.5038/1944-0472.14.2.1910

Wells, G., Horwitz, J., & Seetharaman, D. (2021, September 14). *Facebook knows Instagram is toxic for teen girls, company documents show*. The Wall Street Journal. Retrieved

November 13, 2022, from

<https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739>

Witt, T. (2020). 'If i cannot have it, i will do everything i can to destroy it.' the canonization of Elliot Rodger: 'Incel' masculinities, secular sainthood, and justifications of ideological violence. *Social Identities, 26*(5), 675-689. doi: 10.1080/13504630.2020.1787132

Valenzuela, S., Piña, M., & Ramírez, J. (2017). Behavioral effects of framing on social media users: How conflict, economic, human interest, and morality frames drive news sharing. *Journal of communication, 67*(5), 803-826.

Ziems, C., He, B., Soni, S., & Kumar, S. (2020). Racism is a virus: Anti-asian hate and counterhate in social media during the covid-19 crisis. arXiv preprint arXiv:2005.12423