Samuel Hansen, Sukolsak Sakshuwong, Julie Zhu
December 8, 2015
MS&E 231: Introduction to Computational Social Science

Deindividuation Theory in Anonymous Online Groups

Introduction

In the aftermath of World War II, psychologists Festinger, Pepitone, and Newcomb (1952) hypothesized that "when group members are not seen as individuals, a state of deindividuation may result, with a consequent lowering of restraints" (Diener et al., 1976). This theory sparked a spate of psychological investigation concerning deindividuation in groups, seeking to find "what forces crowd members ... to behave in uncivilized and violent ways" (Diener & Wallbom, 1976).

Early theories of deindividuation attributed "loss of individuality [to] submergence in the crowd" (Postemes & Spears, 1998), suggesting that a variety of group properties such as size and cohesion may affect antinormative behavior. In Diener et al.'s study, *Effects of Deindividuation Variables on Stealing Among Halloween Trick-or-Treaters*, the authors demonstrated that the interaction between group size and anonymity produced the highest stealing rates of candy from bowls labeled "take one only" (Diener et al., 1976). In turn, the researchers suggested self-restraint decreases in larger anonymized groups.

According to deindividuation theory, larger groups afford more anonymity to individual members, resulting in more negative behaviors (Zimbardo, 1969). This hypothesis has been tested in controlled laboratory settings (Kugihara, 2001) and confirmed by a meta-analysis of 61 deindividuation publications (Postemes & Spears, 1998), which found that group size is a significant predictor of antinormative behavior.

Although it is valuable to study the effects of deindividuation in laboratory settings, the recent emergence of online social networks offers new opportunities to test theories of deindividuation at an aggregate scale. Indeed, web forums offering anonymous or pseudo-anonymous interactions are natural laboratories to examine massive group behavior.

A modern offshoot of deindividuation theory is the online disinhibition effect (ODE), coined by cyberpsychologist John Suler as the tendency of online users to "act out more frequently or intensely than they would in person" (2004). Disinhibited online behavior has been attributed to a variety of instances of cyber-vigilantism, including rampant racism in the anonymous message forum 4chan (Siegel, 2015), "fat-shaming" on Reddit.com (Dewey, 2015), and cascades of negative Yelp reviews against Dr. Walter J. Palmer, a big-game hunter who killed a protected lion named Cecil in Zimbabwe (Capecchi and Rogers, 2015). Although such instances offer support for the ODE, little is understood about the underlying factors that mitigate or facilitate anonymous online herd behavior. In turn, we sought to investigate the properties of anonymous online groups that affect negative behavior, testing how traditional theories of deindividuation apply to modern internet settings.

Research Question

This project investigates the online disinhibition effect on the web forum Reddit.com, the self-proclaimed "Front Page of the Internet." Founded in 2005, Reddit allows users, called "redditors", to post content in the form of links, pictures, and text to themed subgroups called "subreddits". As of December 8, 2015, there were 754,920 subreddits registered on the site, ranging from r/funny, a subreddit for humorous content, to r/Showerthoughts, a subreddit "to share anything that goes on in your head whilst in the shower" (Top Subreddits). On each post, redditors can leave comments that are "upvoted" or "downvoted" by other users. A user's "karma score" is determined by the number of upvotes minus downvotes they receive on a post or comment. In turn, Reddit is an ecosystem of pseudo-anonymous[1] individuals posting, commenting, and voting within groups. We chose to study Reddit because groups are predefined as subreddits, and user data is publically accessible via Reddit's API.

In examining this ecosystem, we sought to identify attributes of anonymous groups that affect disinhibited negative behavior. More specifically, we investigated the effects of group size, diversity, and quality on sentiment derived from comments. Drawing from deindividuation literature, we hypothesized that negative sentiment would increase as group anonymity increases. Thus, we would expect group size to increase anonymity, resulting in more negative sentiment in a given subreddit. In the same vein, we would expect greater user diversity to increase deindividuation, and in turn, negative sentiment. Lastly, we tested the effect of the proportion of high-quality users in a group - defined as those with high karma-to-comment ratios - on comment sentiment. Our investigation of this variable was not grounded in theory due to the absence of literature concerning high-quality users and online sentiment; however, we hypothesized that high-quality users may steer conversations amongst subreddits, with possible implications for antinormative behavior.

In testing the effects of group size, diversity, and proportion of quality users on comment sentiment, we seek to compare theories of deindividuation derived from laboratory settings to new evidence obtained from a large anonymous online forum.

Methods

Data Collection

Our primary data source is *Reddit's Comment Dataset*[2], a corpus of 50 million comments spanning January 2008 to December 2012 obtained from Reddit's API (5 GB compressed; 30 GB uncompressed). This data set was featured on Kaggle.com, a data science website, and was originally compiled by Reddit user Jason Baumgartner (Singer et al., 2014). The file consists of JSON objects delimited by new lines (\n). Each JSON object represents a comment, including information such as author, time created, upvotes, downvotes, score (the difference between upvotes and downvotes), subreddit label, and comment text body. Since the comment file is extremely large, we only examined comments from the top 300 subreddits, as ranked by the total number of comments in each subreddit.

Data Cleaning

Since JSON objects included extraneous information (e.g. link ID, time retrieved, user flair, etc.), we wrote a Python script to only select the following variables: subreddit, subreddit_id, link_id, name, parent_id, body, author, created_utc, ups, downs, and score. Furthermore, we only considered users with at least 5 comments to control for outlier effects (e.g. one user posting a single comment that receives 1,000 upvotes). We chose 5 comments as a threshold because, upon inspection of the data, it provided the best tradeoff between the bias introduced by a high threshold (e.g. 10 or 15) and the variance produced by a low one (e.g. 2 or 3). We did not include comments from users with deleted accounts because such

[1] User IDs are public but do not reveal personal details.
[2] https://www.reddit.com/r/datasets/comments/3bxlg7/

missing data impeded computation of aggregate statistics. This yielded a final sample of 36,474,025 comments from 300 unique subreddits.

To identify "high-quality" users, we computed the ratio of each user's karma (upvotes minus downvotes) to the total number of their comments in a given subreddit. The 90th percentile of average karma per comment across all subreddits was approximately 12.8929. We classified users at or above this threshold as "high-quality" and those below it as "regular".

Sentiment Analysis

To derive sentiment from Reddit comments, we used the *Language Inquiry and Word Count* (LIWC)[3] system, a sentiment analysis tool that analyzes text and calculates the percentage of words that are positive or negative (among other text variables). LIWC is a popular sentiment classifier for psychological studies because "subjective dictionaries were independently rated by judges" (Pennebaker and King, 1999). LIWC uses a "bag of words" model that classifies words individually by their semantic category (e.g. positive emotion, negative emotion) and does not consider grammatical structure.

For each comment, the LIWC divided the number of positive and negative emotion words by the total number of words to obtain percentages of words in each affective category. For instance, the sentence "The dog seemed happy" would receive a positive emotion score of 0.25 because there is one positive emotion word ("happy") and four total words, yielding $1 \div 4 = 0.25$. There are no negative words, which are pre-categorized by independent judges, so the negative emotion score is $0 \div 4 = 0$.

LIWC has been used in many studies of internet-based text, including microloan webpages on Kiva.org (Genevsky and Knutson, 2014), messages on online cancer support groups (Lieberman and Goldstein, 2006), and internet depression forums (Ramirez-Esparza et al., 2008). We chose LIWC because of its established use in many related psychological studies.

To classify comment sentiment, we first split our data set of approximately 37 million comments into 37 segments of 1 million comments to meet the computational capacities of the LIWC system. We uploaded each file into LIWC, which returned an output file with appended columns corresponding to the positive and negative emotion scores of each comment. The output file also included variables such as percent of adjectives, percent of pronouns, and total word count; however, we did not consider these variables in our analysis.

Computing Summary Statistics

Using the 37 output files generated by LIWC, we performed a local MapReduce job to join together the files, then separate them into 300 different files containing the data of each subreddit. We wrote a script to iterate through each subreddit's data to compute its aggregate statistics (e.g. number of comments, number of users, average sentiment scores). The average positive and negative sentiment scores were computed for each user in each subreddit by averaging the scores of each user, then averaging these user averages into a single positive or negative score for a given subreddit. The aggregated statistics were written to a final output file with 300 rows, one corresponding to each subreddit.

[3] http://liwc.wpengine.com

Results

Distribution of Group Size

In our preliminary analysis of subreddit-level data, we first examined the distributions of group size as defined by a) number of comments, and b) number of unique users. We used both metrics so that our analysis was robust to more than one definition of group size. Histograms in Figure 1 reveal that both metrics are right-skewed; however, the distribution of number of users appears slightly more normal. Note the differences in scale on the x-axis: the number of comments is an order of magnitude larger than the number of users (since each user commented at least 5 times).

To test for normality of the distributions, we performed Shapiro-Wilk normality tests and were able to reject the null hypothesis that the number of comments and number of users are normally distributed across subreddits ($p < 0.01$) at a 0.05 significance level. Despite slight differences in skew, these metrics tell the same story: the top 300 subreddits are not normally distributed in group size. In turn, we bear in mind that patterns we observe may be due to the fact that most subreddit sizes are centered around the lower end of the scale rather than uniformly distributed across a range. In the top 300 subreddits, the number of comments ranges from 27,091 to 4,712,795 and the number of users ranges from 55 to 606,169, allowing for comparisons of group size.

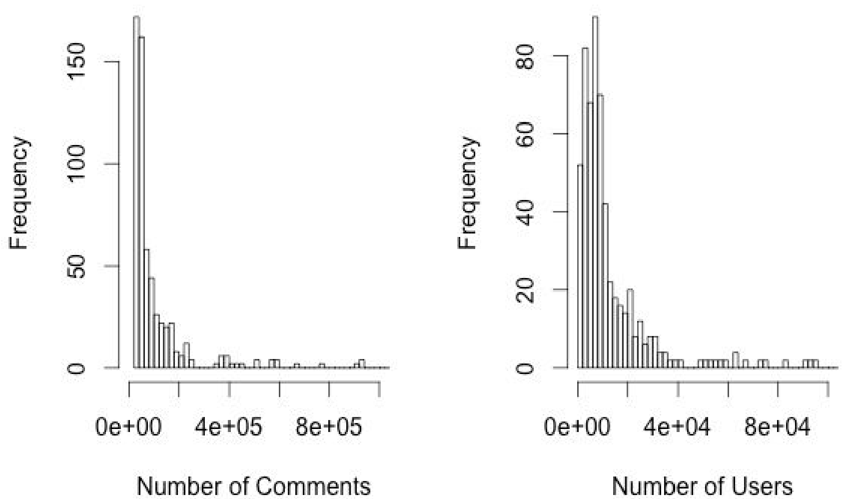

Fig. 1: Subreddit distributions of group size measured by (a) number of comments, and (b) number of unique users.

Distribution of Sentiment

An analysis of average sentiment across subreddits reveals two distinct patterns: first, positive sentiment is higher on average than negative sentiment. As shown in Figure 2b, average positive sentiment is centered around 5.12, while average negative sentiment is centered around 2.39. Second, a Shapiro-Wilk test of normality revealed that negative sentiment is normally distributed across subreddits (p-value of 0.34, using a significance level of 0.05), whereas positive sentiment is not (p-value < 0.001). This is confirmed by the apparent right skew of positive sentiment in Figure 2b. We see that even though positive sentiment is heavily right skewed, negative sentiment is more normally distributed. Taken together, these results raise the question, what underlying mechanisms give rise to differences in comment sentiment?

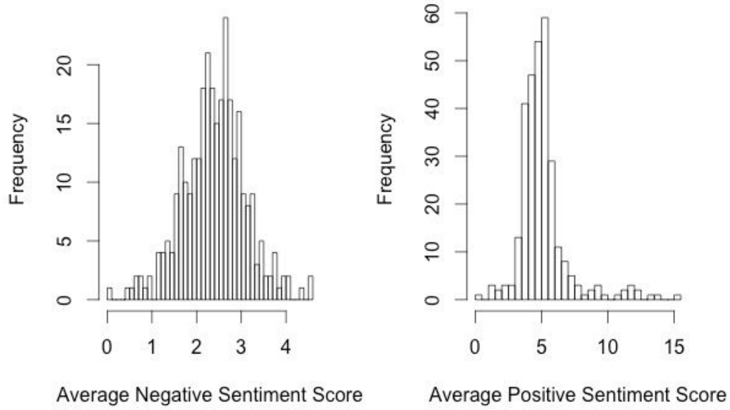

Fig. 2: Subreddit distributions of (a) average negative sentiment score, and (b) average positive sentiment score.

Group Size vs. Sentiment

Drawing from deindividuation literature, we first tested the effect of group size on comment sentiment. Again, deindividuation theory predicts larger groups afford more anonymity to individual members, resulting in more deindividuation and negative behaviors. We observed this precise effect using both metrics of group size: a) number of comments, and b) number of users. Standard Pearson correlation analyses indicated negative sentiment has a strong, positive relationship with both number of comments (0.15) and users (0.14), and that positive sentiment has slightly weaker and negative relationship with comments (-0.05) and users (-0.11).

To estimate the significance of these relationships, we fit linear models regressing average negative sentiment onto log of number of comments (Fig. 3a) and log of number of users (Fig. 3b), and found significant ($p < 0.01$) positive coefficients in both models (see Table 1).

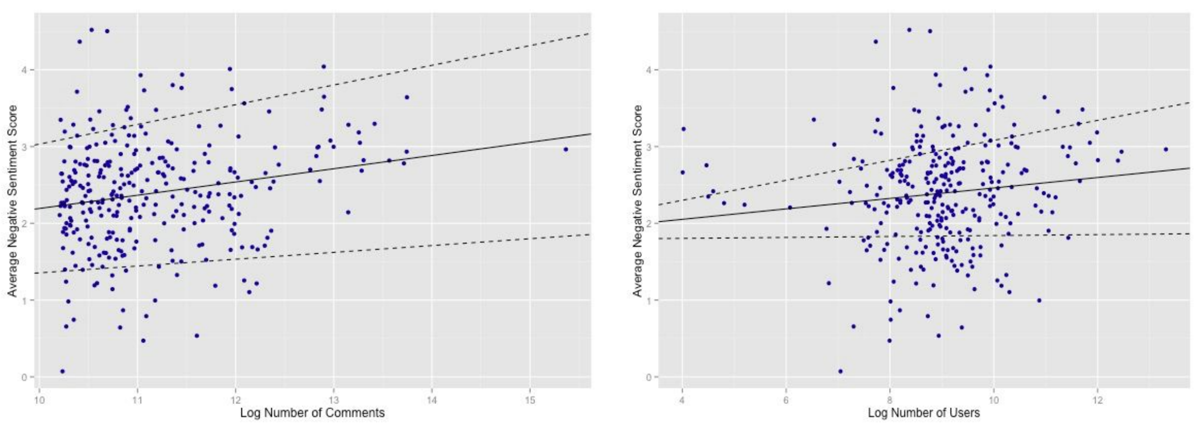

Fig 3: Scatterplots of subreddit group size measured by (a) log number of comments and (b) log number of users versus average negative sentiment. Dark black lines represent coefficients of linear model fits, and dashed lines represent their standard errors.

We repeated the same analysis for positive sentiment, but obtained inconsistent results: positive sentiment has a significant negative relationship with number of users ($p < 0.01$), but not with number of comments (p > 0.1), where we in fact obtained a slightly positive and near zero coefficient of 0.008 (see Table 1). This inconsistency yields a non-robust result for positive sentiment and is addressed further in the discussion section.

	Dependent variable:			
	Positive Sentiment	Negative Sentiment	Positive Sentiment	Negative Sentiment
	(1)	(2)	(3)	(4)
Log Number of Comments	0.008	0.200***		
	(0.141)	(0.049)		
Log Number of Users			-0.388***	0.128***
			(0.089)	(0.032)
Constant	5.481***	0.298	9.070***	1.376***
	(1.582)	(0.548)	(0.815)	(0.291)
Observations	300	300	300	300
R2	0.00001	0.053	0.059	0.051
Adjusted R2	-0.003	0.050	0.056	0.048
Residual Std. Error (df = 298)	2.051	0.711	1.989	0.711
F Statistic (df = 1; 298)	0.003	16.722***	18.847***	16.061***
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 1: Summary of coefficients from models regressing average positive and negative sentiment onto log number of comments and log number of users. Negative sentiment has a significant positive relationship with log number of comments and users, whereas positive sentiment has a significant negative relationship with log number of users, but not comments.

Group Quality vs. Sentiment

Subsequently, we tested the effect of group quality on comment sentiment. Users are considered to be "high-quality" if they are in the 90th percentile of average karma per comment. Practically speaking, these users receive the most upvotes for each comment they post. In turn, we measured group quality by the percentage of high-quality users in a given subreddit. A Pearson correlation analysis yielded a clear result: negative sentiment is positively correlated with group quality (0.41)**,** whereas positive sentiment is negatively correlated with it (-0.29). Linear regressions of negative sentiment and positive sentiment onto percent of quality users confirmed these relationships (see Table 2), which are plotted in Figure 4. Overall, this implies negative sentiment increases as group quality improves (see discussion section for interpretation of results).

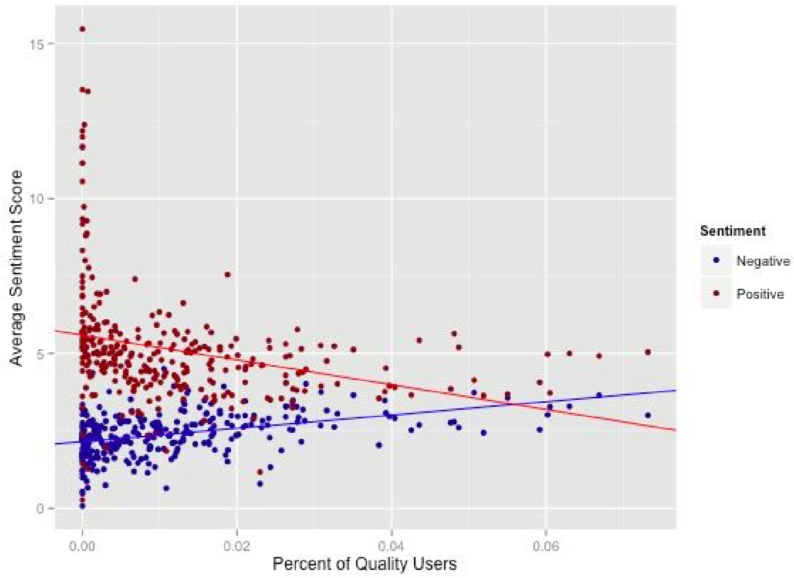

Fig 4: Percentages of quality users in a subreddit versus comment sentiment scores. Red and blue lines represent linear regression fits for positive and negative sentiment, respectively. Each point represents a single subreddit.

	Dependent variable:	
	Positive Sentiment (1)	Negative Sentiment (2)
Proportion of Quality Users	-43.450*** (8.365)	
Proportion of Quality Users		25.102*** (2.749)
Constant	6.068*** (0.149)	2.243*** (0.049)
Observations	300	300
R2	0.083	0.219
Adjusted R2	0.080	0.216
Residual Std. Error (df = 298)	1.964	0.646
F Statistic (df = 1; 298)	26.980***	83.360***
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 2: Summary of coefficients from models regressing average positive and negative sentiment onto proportion of quality users. Negative sentiment has a significant positive relationship with proportion of quality users, whereas positive sentiment has a significant negative relationship.

Group Diversity vs. Sentiment

To measure the diversity of a subreddit, we looked for a diversity index that measures both richness and equality of comments in the subreddit. Richness means that, given two subreddits with the same number of comments, the subreddit with more users should be more diverse than the subreddit with fewer users. Equality means that, given two subreddits with the same number of users, the subreddit whose users made an equal number of comments should be more diverse than the subreddit whose comments are disproportionately made by a particular group of users.

The Shannon index, a diversity index that measures the entropy of data, satisfies both of these requirements. It is defined as $-\Sigma_{i=1}^{n} p_i \, ln \, p_i$, where p_i is the proportion of data belonging to the *i*th type and *n* is the total number of types. In our case, p_i is the ratio of the number of comments by user *i* to the total number of comments, and *n* is the total number of users. Using the Shannon index, we found small correlations between diversity and positive sentiment (-0.19) and negative sentiment (0.19). Coefficient estimates from linear regressions were small, yet significant for both positive (-0.331) and negative (0.15) sentiment ($p < 0.01$). Taken together, these results imply negative sentiment increases as group diversity increases, whereas positive sentiment decreases.

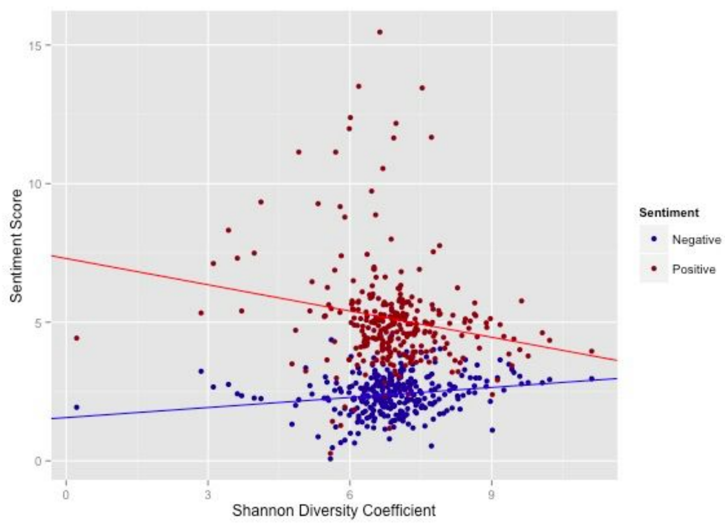

Fig. 5: Subreddit Shannon diversity coefficients versus comment sentiment scores. Red and blue lines represent linear regression fits for positive and negative sentiment, respectively. Each point represents a single subreddit.

To test whether these results were robust across different diversity measures, we repeated our analysis using the Gini-Simpson index, which measures the probability that two randomly drawn comments come from different authors. Mathematically, the Gini-Simpson index is defined as 1-λ, where λ equals $\Sigma_{i=1}^{n} p_i^2$, the probability that two entities taken at random from the dataset of interest (with replacement) represent the same type. In our specific data set, we define a "type" to be a user. Thus, when the index is small, there is a small probability that two randomly drawn comments come from different authors. This corresponds to a low-diversity group is which a few users post all the comments. Conversely, when the Gini-Simpson index is large, there is a large probability that two randomly drawn comments come from the differents authors, corresponding to a high diversity group in which all members post equal number of comments.

Interestingly, linear regressions of the Gini-Simpson diversity index onto positive and negative sentiment revealed small and insignificant coefficients ($p > 0.1$) for positive (0.622) and negative sentiment (-0.376). While consistent with our results using the Shannon index, the results obtained from using the Gini-Simpson diversity index suggested that group diversity has no significant effect on sentiment. Because both diversity indices yield small coefficients, and in one case, insignificant results, we cannot discern a clear pattern between group diversity and sentiment.

Discussion

Group Size vs. Sentiment

Based on our results, we conclude the following: first, negative sentiment increases with group size. These results are consistent with those predicted by the online disinhibition effect: as groups become bigger, individuals experience more anonymity, and demonstrate more disinhibited negative behaviors. This effect was robust across both measures of group size: number of users and comments. Although positive sentiment did not reveal consistent results across both measures of group size, its significant negative relationship with log number of users supports the pattern we observe concerning negative sentiment: larger groups are associated with more negativity. The non-robust result for positive sentiment may be attributable to its non-normal distribution across subreddits (see Figure 2b).

In summary, when groups are defined by subreddits and size is defined by number of users or comments, we observe the same positive relationship between group size and negative behavior shown by decades of psychological research concerning deindividuation theory (Diener et al., 1976; Postemes & Spears, 1998; Kugihara, 2001, etc.). In effect, we were able to replicate a key component of deindividuation theory in the context of anonymous groups on the internet.

Group Quality vs. Sentiment

Second, we conclude that negative sentiment increases with group quality, whereas positive sentiment decreases with it. This result may imply a group "ringleader effect", in which high-quality users – those who post consistently popular comments – rally negativity from other users. It is possible high-quality users wield more influence in comment threads, meaning a comment by a high-quality user may ignite a chain-reaction of negativity from other users. This invites further examination concerning the distribution of sentiment across and within subreddits to investigate whether negativity is normally distributed or caused by a few "bad apples".

In an initial attempt to answer this question, we hypothesized that if negative sentiment in a subreddit follows a normal distribution, then its users are contributing evenly to overall negativity. On the other hand, we hypothesized that if the distribution of negative sentiment is heavily skewed in a subreddit, then a few "bad apple" users with particularly negative sentiment scores are the main contributors.

In an initial comparison of negative sentiment between subreddits, we randomly selected two subreddits with average negative sentiment scores from the 75th percentile and the 25th percentile (r/fatpeoplehate and r/askscience, respectively) and plotted the average negative sentiment scores per user (see Figure 6). The subreddit r/fatpeoplehate is "devoted to text and photos that [insult] fat people" (Price, 2015), whereas r/askscience is a forum to "promote scientific literacy" ("AskScience User Help Page").

We observe somewhat different distribution shapes for negative sentiment. In r/fatpeoplehate, a subreddit with high negative sentiment, negative sentiment seems to be more normally distributed (see Figure 6a). However, in r/askscience, a subreddit known to have lower negative sentiment, the sentiment distribution is more right-skewed (see Figure 6b). Indeed, the majority of the sentiment observations are centered on the lower half of the range, suggesting that the negative sentiment for less negative subreddits may be controlled by a few users.

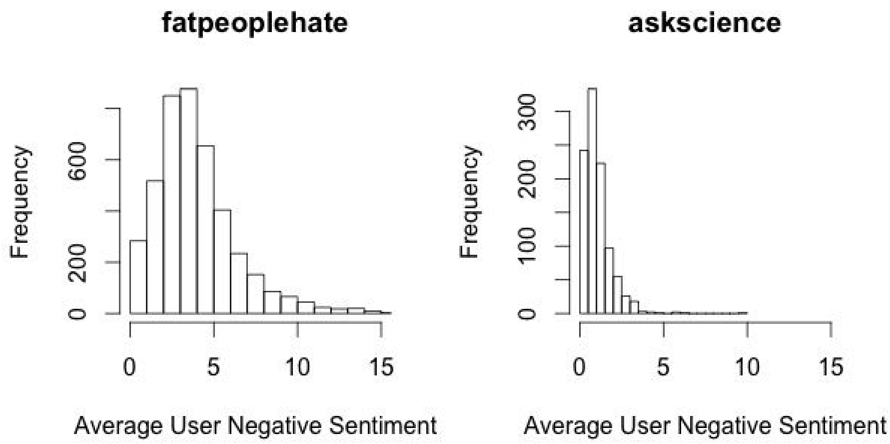

Fig. 6: Distribution of average user negative sentiment for two subreddits: (a) r/fatpeoplehate and (b) r/askscience. r/fatpeoplehate has an average negative sentiment score within the 75th percentile of all 300 subreddits in our study, whereas r/askscience has an average negative sentiment score within the 25th percentile.

Overall, this initial comparison suggests that more negative subreddits have more normally distributed negative sentiment because every user participates in expressing negativity. On the other hand, the distribution of negative sentiment in less negative subreddits may be more skewed due to the general community abstaining from negative sentiment, leaving a few "bad apple" users contributing a disproportionate amount of negativity. More research is needed on this topic, but hopefully these preliminary results will spark further inquiry concerning the contributions of individual users to negative sentiment across subreddits.

Group Diversity vs. Sentiment

Based on deindividuation literature, we hypothesized that greater diversity in a group would afford more anonymity to individual users, increasing deindividuation and, in turn, negative sentiment in a given subreddit. We observed this precise pattern using the Shannon index of diversity: negative sentiment increases with group diversity. This trend was corroborated by using an alternative measure of diversity: the Gini-Simpson index. However, the relationship between the Gini-Simpson index and negative sentiment was not statistically significant. Furthermore, coefficient estimates obtained from linear regressions on both measures of diversity were small, suggesting only a weak relationship between group diversity and negative sentiment. Although these results were consistent with our hypothesis, the weakness of this effect and non-robust results from different measures of diversity lead us to conclude the relationship between group diversity and negative sentiment is inconclusive. Further research is needed to elucidate this relationship.

Source of Error

Definition of Diversity

We computed our two measures of diversity (Shannon index and Gini-Simpson index) using the number of users and comments in a given subreddit. However, a deeper analysis should expand this definition to include other measures of subreddit diversity, including user activity, popularity, location (from IP addresses), and gender (if data are available). Doing so will provide a higher-resolution picture of the demographic diversity factors that influence subreddit sentiment.

LIWC

A major drawback of the LIWC sentiment classifier is its reliance on a “bag-of-words” model, which fails to account for pragmatic meaning, including negation, sarcasm, and long-distance dependencies. This may lead to misclassifying words as having the opposite sentiment than the author intended, which may affect the distributions of positive and negative sentiment across subreddits. A sentiment analysis tool that accounts for syntactic structure, such as the *Recursive Neural Tensor Network* (Socher et al., 2013), may yield more accurate classifications of comment sentiment.

Further Research

Our research represents a quantitative analysis of the factors affecting subreddit sentiment; however, our findings would benefit from an assessment of the qualitative features of subreddits that influence comment sentiment as well. For instance, we manually inspected the 20 most negative subreddits in the 75th percentile of average negative sentiment and found most were devoted to themes including sports (r/GreenBayPackers, r/cowboys, r/CollegeBasketball), race-related content (r/BlackPeopleTwitter, r/4chan), or miscellaneous negative themes (r/ImGoingToHellForThis, r/fatpeoplehate, r/JusticePorn). By contrast, the 20 least negative subreddits (in the 25th percentile of average negative sentiment) included themes such as “do-it-yourself” activities (r/HomeBrewing, r/woodworking, r/buildapc) and giving advice (r/weddingplanning, r/personalfinance). In this way, the qualitative theme of a subreddit can be informative about its corresponding level of negativity.

Further research should investigate how the qualitative feature of subreddit theme interacts with quantitative variables such as group size, quality, and diversity. Doing so will provide a more holistic understanding of online group behavior. By addressing theories of deindividuation at the online scale, we hope our findings contribute to the ongoing investigation of human behavior in anonymous groups.

Works Cited

"AskScience User Help Page." *Reddit.com*. N.p., 2015. Web.

Capecchi, Christina, and Katie Rogers. "Killer of Cecil the Lion Finds Out That He Is a Target Now, of Internet Vigilantism." The New York Times. The New York Times, 29 July 2015. Web. 15 Oct. 2015.

Dewey, Caitlin. "Censorship, Fat-shaming and the 'Reddit Revolt': How Reddit Became the Alamo of the Internet's Ongoing Culture War." Washington Post. The Washington Post, 12 June 2015. Web. 15 Oct. 2015.

Diener, Edward et al. "Effects of Deindividuation Variables on Stealing among Halloween Trick-or-Treaters." *Journal of Personality and Social Psychology* 33.2 (1976): 178–183. Web.

Diener, Edward, and Mark Wallbom. "Effects of Self-Awareness on Antinormative Behavior." *Journal of Research in Personality* 10.1 (1976): 107–111. Web.

Festinger, L. Pepitone, A. and Newcomb T. (1952). Some Consequences of Deindividuation in a Group. *Journal of Abnormal and Social Psychology,* 47, 382-389.

Genevsky, A., and B. Knutson. "Neural Affective Mechanisms Predict Market-Level Microlending." *Psychological Science* 26.9 (2015): 1411–1422. Web.

Kugihara, N. "Effects of Aggressive Behaviour and Group Size on Collective Escape in an Emergency: A Test between a Social Identity Model and Deindividuation Theory." *The British journal of social psychology / the British Psychological Society* 40.Pt 4 (2001): 575–598. Web.

Lieberman, Morton A., and and Benjamin A. Goldstein. "Not All Negative Emotions Are Equal: The Role of Emotional Expression in Online Support Groups for Women with Breast Cancer." *Psycho-Oncology* 15.May 2005 (2006): 160–168. Web.

Pennebaker, James W., and Laura A. King. "Linguistic Styles: Language Use as an Individual Difference." *Journal of Personality and Social Psychology* 77.1296 (1999): Print.

Postmes, Tom, and Russell Spears. "Deindividuation and Antinormative Behavior: A Meta-Analysis." *Psychological Bulletin* 123.3 (1998): 238–259. Web.

Price, Rob. "Reddit Users Are Furious after the Site Banned a Community Dedicated to Hating Fat People." *Business Insider* 2015. Web.

Ramirez-Esparza, N et al. “The Psychology of Word Use in Depression Forums in English and in Spanish: Testing Two Text Analytic Approaches.” ... *Conference on Weblogs ...* (2008): 102–108. Web.

Siegel, Jacob. "Dylann Roof, 4chan, and the New Online Racism." The Daily Beast. Newsweek/Daily Beast, 29 June 2015. Web. 15 Oct. 2015.

Singer, Philipp et al. “Evolution of Reddit: From the Front Page of the Internet to a Self-Referential Community ?” *Proceedings of the companion publication of the 23rd international conference on World wide web companion* (2014): 1–6. Web.

Socher, Richard et al. “Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank.” *Conference on Empirical Methods in Natural Language Processing* (2013): n. pag. Print.

Suler, John. "The Online Disinhibition Effect." CyberPsychology & Behavior 7.3 (2004): 321-26. Web.

Zimbardo, P. G. (1970). The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. In W. J. Arnold & D. Levine (Eds.), 1969 Nebraska Symposium on Motivation (pp. 237-307). Lincoln, NE: University of Nebraska Press.