Identifying Dogmatism in Social Media: Signals and Models

Ethan Fast and **Eric Horvitz**
`ethaen@stanford.edu, horvitz@microsoft.com`

Abstract

We explore linguistic and behavioral features of dogmatism in social media and construct statistical models that can identify dogmatic comments. Our model is based on a corpus of Reddit posts, collected across a diverse set of conversational topics and annotated via paid crowdsourcing. We operationalize key aspects of dogmatism described by existing psychology theories (such as over-confidence), finding they have predictive power. We also find evidence for new signals of dogmatism, such as the tendency of dogmatic posts to refrain from signaling cognitive processes. When we use our predictive model to analyze millions of other Reddit posts, we find evidence that suggests dogmatism is a deeper personality trait, present for dogmatic users across many different domains, and that users who engage on dogmatic comments tend to show increases in dogmatic posts themselves.

1 Introduction

> "I'm supposed to trust the opinion of a MS minion? The people that produced Windows ME, Vista and 8? They don't even understand people, yet they think they can predict the behavior of new, self-guiding AI?" *–anonymous*

> "I think an AI would make it easier for Patients to confide their information because by nature, a robot cannot judge them. Win-win? :D"' *–anonymous*

Dogmatism describes the tendency to lay down opinions as incontrovertibly true, without respect for conflicting evidence or the opinions of others (Oxford Dictionary, 2016). Which user is more dogmatic in the examples above? This question is simple for humans. Phrases like "they think" and "they don't even understand," suggest an intractability of opinion, while "I think" and "win-win?" suggest the opposite. Can we train computers to draw similar distinctions? Work in psychology has called out many aspects of dogmatism that can be modeled computationally via natural language, such as over-confidence and strong emotions (Rokeach, 1954).

We present a statistical model of dogmatism that addresses two complementary goals. First, we validate psychological theories by examining the predictive power of feature sets that guide the model's predictions. For example, do linguistic signals of certainty help to predict a post is dogmatic, as theory would suggest? Second, we apply our model to answer four questions:

**R1**: What kinds of topics (e.g., guns, LGBT) attract the highest levels of dogmatism?

R2: How do dogmatic beliefs cluster?

R3: How does dogmatism influence a conversation on social media?

R4: How do other user behaviors (e.g., frequency and breadth of posts) relate to dogmatism?

We train a predictive model to classify dogmatic posts from Reddit, one of the most popular discussion communities on the web.[1] Posts on Reddit capture discussion and debate across a diverse set of domains and topics – users talk about everything from climate change and abortion, to world news and relationship advice, to the future of artificial intelligence. As a prerequisite to training our model, we have created a corpus of 5,000 Reddit posts annotated with levels of dogmatism, which we are releasing to share with other researchers.

---

[1] http://www.reddit.com

690
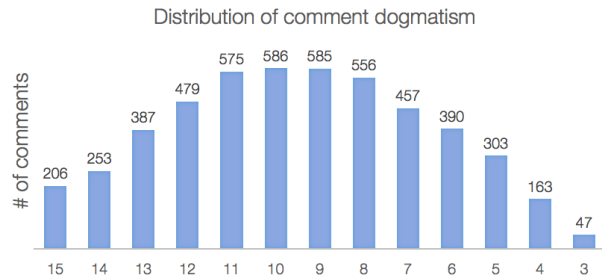
Figure 1: We crowdsourced dogmatism labels for 5000 comments. The distribution is slightly skewed towards higher levels of dogmatism. For example, crowdworkers unanimously labeled 206 comments as highly dogmatic ($5 \times 3 = 15$), but only 47 as minimally dogmatic ($1 \times 3 = 3$).

Using the model, we operationalize key domain-independent aspects of psychological theories of dogmatism drawn from the literature. We find these features have predictive power that largely supports the underlying theory. For example, posts that use less confident language tend to be less dogmatic. We also discover evidence for new attributes of dogmatism. For example, dogmatic posts tend not to verbalize cognition, through terms such as "I think," "possibly," or "might be."

Our model is trained on only 5,000 annotated posts, but once trained, we use it to analyze millions of other Reddit posts to answer our research questions. We find a diverse set of topics are colored by dogmatic language (e.g., people are dogmatic about religion, but also about LGBT issues). Further, we find some evidence for dogmatism as a deeper personality trait – people who are strongly dogmatic about one topic are more likely to express dogmatic views about others as well. Finally, in conversation, we discover that one user's dogmatism tends to bring out dogmatism in their conversational partner, forming a vicious cycle.

2 Dogmatism data

Posts on Reddit capture debate and discussion across a diverse set of topics, making them a natural starting point for untangling domain-independent linguistic features of dogmatism.

Data collection. Subreddits are sub-communities on Reddit oriented around specific interests or topics, such as *technology* or *politics*. Sampling from Reddit as a whole would bias the model towards the most commonly discussed content. But by sampling posts from individual subreddits, we can control the kinds of posts we use to train our model. To collect a diverse training dataset, we have randomly sampled 1000 posts each from the subreddits *politics*, *business*, *science*, and *AskReddit*, and 1000 additional posts from the Reddit frontpage. All posts in our sample appeared between January 2007 and March 2015, and to control for length effects, contain between 300 and 400 characters. This results in a total training dataset of 5000 posts.

Dogmatism annotations. Building a useful computational model requires labeled training data. We labeled the Reddit dataset using crowdworkers on Amazon Mechanical Turk (AMT), creating the first public corpus annotated with levels of dogmatism. We asked crowdworkers to rate levels of dogmatism on a 5-point Likert scale, as supported by similar annotation tasks in prior work (Danescu-Niculescu-Mizil et al., 2013). Concretely, we gave crowdworkers the following task:

> Given a comment, imagine you hold a well-informed, different opinion from the commenter in question. We'd like you to tell us how likely that commenter would be to engage you in a constructive conversation about your disagreement, where you each are able to explore the other's beliefs. The options are:
>
> **(5)**: It's unlikely you'll be able to engage in any substantive conversation. When you respectfully express your disagreement, they are likely to ignore you or insult you or otherwise lower the level of discourse.
>
> **(4)**: They are deeply rooted in their opinion, but you are able to exchange your views without the conversation degenerating too much.
>
> **(3)**: It's not likely you'll be able to change their mind, but you're easily able to talk and understand each other's point of view.
>
> **(2)**: They may have a clear opinion about the subject, but would likely be open to discussing alternative viewpoints.
>
> **(1)**: They are not set in their opinion, and it's possible you might change their mind. If the comment does not convey an opinion of any kind, you may also select this option.

To ensure quality work, we restricted the task to Masters workers and provided examples corresponding to each point on the scale. Including examples in a task has been shown to significantly increase the agreement and quality of crowdwork

(Doroudi et al., 2016). For instance, here is an example of a highly dogmatic *(5)* comment:

> I won't be happy until I see the executive suite of BofA, Wells, and all the others, frogmarched into waiting squad cars. It's ALREADY BEEN ESTABLISHED that...

And a minimally dogmatic *(1)* comment:

> I agree. I would like to compile a playlist for us trance yogi's, even if you just would like to experiment with it. Is there any preference on which platform to use?

Each comment has been annotated by three independent workers on AMT, which is enough to produce reliable results in most labeling tasks (Sheng et al., 2008). To compute an aggregate measure of dogmatism for each comment, we summed the scores of all three workers. We show the resulting distribution of annotations in Figure 1.

Inter-annotator agreement. To evaluate the reliability of annotations we compute Krippendorff's α, a measure of agreement designed for variable levels of measurement such as a Likert scale (Hayes and Krippendorff, 2007). An α of 0 indicates agreement indistinguishable from chance, while an α of 1 indicates perfect agreement. Across all annotations we find $\alpha = 0.44$. While workers agree much more than chance, clearly dogmatism is also subjective. In fact, when we examine only the middle two quartiles of the dogmatism annotations, we find agreement is no better than chance. Alternatively, when we measure agreement only among the top and bottom quartiles of annotations, we find agreement of $\alpha = 0.69$. This suggests comments with scores that are only slightly dogmatic are unreliable and often subject to human disagreement. For this reason, we use only the top and bottom quartiles of comments when training our model.

3   Approaches to Identifying Dogmatism

We now consider strategies for identifying dogmatism based on prior work in psychology. We start with the Linguistic Inquiry and Word Count (LIWC), a lexicon popular in the social sciences (Pennebaker et al., 2001). LIWC provides human validated lists of words that correspond to high-level psychological categories such as *certainty* or *perception*. In other studies, LIWC has uncovered

linguistic signals relating to politeness (Danescu-Niculescu-Mizil et al., 2013), deception (Yoo and Gretzel, 2009), or authority in texts (Gilbert, 2012). Here, we examine how dogmatism relates to 17 of LIWC's categories (Table 1).

To compute the relationships between LIWC categories and dogmatism, we first count the relevant category terms that appear in each annotated Reddit comment, normalized by its word count. We then calculate odds ratios on the aggregate counts of each LIWC category over the top and bottom quartiles of dogmatic comments. As we have discussed, using the top and bottom quartiles of comments provides a more reliable signal of dogmatism. We check for significant differences in categories between dogmatic and non-dogmatic comments using the Mann-Whitney U test and apply Holmes method for correction. All odds we report in this section are significant after correction.

Dogmatic statements tend to express a high degree of certainty (Rokeach, 1954). Here we consider LIWC categories that express certainty both positively (*certainty*) and negatively (*tentativeness*). For example, the word "always" is certain, while "possibly" is tentative. Conforming to existing theory, *certainty* is more associated with dogmatic comments (1.52 odds), while *tentativeness* is more associated with the absence of dogmatism (0.88 odds).

Terms used to verbalize cognition can act as a hedge that often characterizes non-dogmatic language. LIWC's *insight* category captures this effect through words such as "think," "know," or "believe." These words add nuance to a statement (Pennebaker and Francis, 1996), signaling it is the product of someone's mind ("*I think* you should give this paper a good review") and not meant to be interpreted as an objective truth. Along these lines, we find the use of terms in the *insight* category is associated with non-dogmatic comments (0.83 odds).

Sensory language, with its focus on description and detail, often signals a lack of any kind of opinion, dogmatic or otherwise. LIWC's *perception* category captures this idea through words associated with hearing, feeling, or seeing. For example, these words might occur when recounting a personal experience ("I saw his incoming fist"), which even if emotionally charged or negative, is less likely to be dogmatic. We find perception is associated with

Strategy	Odds	Example
Certainty	1.33*	Be a hate monger **all** you want... Your life will **never** truly be happy though, and you will **never** know peace.
Tentativeness	0.88*	Most are **likely** to be more technically advanced and, **if** still using radio, **might** very well be emitting signals we could detect
Insight	0.83*	I **think** stating the obvious is a necessary function. Information like this is important to **consider**...
Perception	0.77*	I **saw** four crows on that same branch, **staring** at the deceased. The **silence** of the crows was **deafening**.
Relativity	0.82*	I've known a number to go **into** shock **during** the procedure
Comparison	0.91	This may be **more than** a coincidence.
I (pronouns)	0.68*	Like **I** said, **I** want to believe the former. **I'm** glad it worked out.
You (pronouns)	2.18*	I don't give a fuck what **you** do. **You** can get drink **yourself** to death, **you** can get **yourself** pregnant...
We (pronouns)	0.96	**We** need a bigger, better, colder fridge. **We** have worked hard...
They (pronouns)	1.63*	**They** want the ability to prosecute who they please.
Past	0.69*	I **was** walking past and **thought** about asking if they **needed** help.
Present	1.11*	Can I **steal** your organs and nutrients if I **need** them and you don't **want** to give them up?
Future	1.06	Trump's thugs **will** be pretending to be Bernie supporters and **will** set fire to Philadelphia.
Interrogatory	1.12*	Gee, **where** was the NY Times back in the day? **Why** didn't we hear of the Kennedys, LBJ and FDR?
Negation	1.35*	If you **didn't** know the woman well enough to know she **didn't** take BC regularly, you certainly **don't** know her well enough to know she **doesn't** have an std.
Negative emotion	2.32*	A prank?!? You **arrogant** son of a **bitch**
Positive emotion	0.96	They were **excellent** fishermen - they built **fine** boats.

Table 1: Linguistic features that capture high level psychological categories and their relationship with dogmatic comments. *Strategy* describes the psychological category. *Odds* describes the likelihood that a category will appear more often in a dogmatic comment (e.g., dogmatic comments are 2.18 times more likely to mention *you*-oriented phrases). *Example* illustrates a comment that matches the category. * indicates significance ($p < 0.05$) after correction with Holmes method.

non-dogmatic comments at 0.77 odds.

Drawing comparisons or qualifying something as relative to something else conveys a nuance that is absent from traditionally dogmatic language. The LIWC categories *comparison* and *relativity* capture these effects through comparison words such as "than" or "as" and qualifying words such as "during" or "when." For example, the statement "I hate politicians" is more dogmatic than "I hate politicians *when* they can't get anything done.' *Relativity* is associated with non-dogmatic comments at 0.80 odds,

but *comparison* does not reach significance.

Pronouns can be surprisingly revealing indicators of language: for example, signaling one's gender or hierarchical status in a conversation (Pennebaker, 2011). We find first person singular pronouns are a useful negative signal for dogmatism (0.46 odds), while second person singular pronouns (2.18 odds) and third person plural (1.63 odds) are a useful positive signal. Looking across the corpus, we see *I* often used with a hedge ("I think" or "I know"), while *you* and *they* tend to characterize the beliefs of oth-

ers, often in a strongly opinionated way ("you are a moron" or "they are keeping us down"). Other pronoun types do not show significant relationships.

Like pronouns, verb tense can reveal subtle signals in language use, such as the tendency of medical inpatients to focus on the past (Wolf et al., 2007). On social media, comments written in the *present tense* are more likely to be oriented towards a user's current interaction ("this *is* all so stupid"), creating opportunities to signal dogmatism. Alternatively, comments in the *past tense* are more likely to refer to outside experiences ("it *was* an awful party"), speaking less to a user's stance towards an ongoing discussion. We find *present tense* is a positive signal for dogmatism (1.11 odds) and *past tense* is a negative signal (0.69 odds).

Dogmatic language can be either positively or negatively charged in sentiment: for example, consider the positive statement "*Trump is the SAVIOR of this country!!!*" or the negative statement "*Are you REALLY that stupid?? Education is the only way out of this horrible mess. It's hard to imagine how anyone could be so deluded.*" In diverse communities, where people hold many different kinds of opinions, dogmatic opinions will often tend to come into conflict with one another (McCluskey and Hmielowski, 2012), producing a greater likelihood of negative sentiment. Perhaps for this reason, *negative emotion* (2.09 odds) is a useful positive signal of dogmatism, while *positive emotion* shows no significant relationship.

Finally, we find that *interrogative* language (1.12 odds) and *negation* (1.35 odds) are two additional positive signals of dogmatism. While interrogative words like "how" or "what" have many benign uses, they disproportionately appear in our data in the form of rhetorical or emotionally charged questions, such as "how can anyone be that dumb?"

Many of these linguistic signals are correlated with each other, suggesting that dogmatism is the cumulative effect of many component relationships. For example, consider the relatively non-dogmatic statement: "I think the reviewers are wrong in this instance." Removing signals of *insight*, we have: "the reviewers are wrong in this instance," which is slightly more dogmatic. Then removing *relativity*, we have: "the reviewers are wrong." And finally, adding *certainty*, we have a dogmatic state-

Classifier	In-domain	Cross-domain
BOW	0.853	0.776
SENT	0.677	0.646
LING	0.801	0.728
BOW + SENT	0.860	0.783
BOW + LING	0.881	0.791

Table 2: The AUC scores for dogmatism classifiers within and across domains. BOW (bag-of-words) and SENT (sentiment signals) are baselines, and LING uses the linguistic features from Table 1. We compute in-domain accuracy using 15-fold cross-validation on the Reddit dataset, and cross-domain accuracy by training on Reddit and evaluating on comments on articles from the New York Times. Chance AUC is 0.5.

ment: "the reviewers are always wrong."

4   Predicting dogmatism

We now show how we can use the linguistic feature sets we have described to build a classifier that predicts dogmatism in comments. A predictive model further validates our feature sets, and also allows us to analyze dogmatism in millions of other Reddit comments in a scalable way, with multiple uses in ongoing, downstream analyses.

Prediction task. Our goal is (1) to understand how well we can use the strategies in Section 3 to predict dogmatism, and (2) to test the *domain-independence* of these strategies. First, we test the performance of our model under cross-validation within the Reddit comment dataset. We then evaluate the Reddit-based model on a held out corpus of New York Times comments annotated using the technique in Section 2. We did not refer to this second dataset during feature construction.

For classification, we consider two classes of comments: *dogmatic* and *non-dogmatic*. As in the prior analysis, we draw these comments from the top and bottom quartiles of the dogmatism distribution. This means the classes are balanced, with 2,500 total comments in the Reddit training data and 500 total comments in the New York Times testing data.

We compare the predictions of logistic regression models based on unigram bag-of-words features (BOW), sentiment signals[2] (SENT), the linguistic

---

[2]For SENT, we use normalized word counts from LIWC's positive and negative emotional categories.

features from our earlier analyses (LING), and combinations of these features. BOW and SENT provide baselines for the task. We compute BOW features using term frequency-inverse document frequency (TF-IDF) and category-based features by normalizing counts for each category by the number of words in each document. The BOW classifiers are trained with regularization (L2 penalties of 1.5).

Classification results. We present classification accuracy in Table 2. BOW shows an AUC of 0.853 within Reddit and 0.776 on the held out New York Times comments. The linguistic features boost classification results within Reddit (0.881) and on the held out New York Times comments (0.791). While linguistic signals by themselves provide strong predictive power (0.801 AUC within domain), sentiment signals are much less predictive.

These results suggest that linguistic features inspired by prior efforts in psychology are useful for predicting dogmatism in practice and generalize across new domains.

5 Dogmatism in the Reddit Community

We now apply our dogmatism classifier to a larger dataset of posts, examining how dogmatic language shapes the Reddit community. Concretely, we apply the BOW+LING model trained on the full Reddit dataset to millions of new unannotated posts, labeling these posts with a probability of dogmatism according to the classifier (0=non-dogmatic, 1=dogmatic). We then use these dogmatism annotations to address four research questions.

5.1 What subreddits have the highest and lowest levels of dogmatism? (R1)

A natural starting point for analyzing dogmatism on Reddit is to examine how it characterizes the site's sub-communities. For example, we might expect to see that subreddits oriented around topics such as abortion or climate change are more dogmatic, and subreddits about cooking are less so.

To answer this question, we randomly sample 1.6 million posts from the entire Reddit community between 2007 and 2015. We then annotate each of these posts with dogmatism using our classifier, and compute the average dogmatism level for each subreddit in the sample with at least 100 posts.

Highest	Score	Lowest	Score
cringepics	0.553	photography	0.399
DebateAChristian	0.551	DIY	0.399
DebateReligion	0.540	homebrewing	0.401
politics	0.536	cigars	0.402
ukpolitics	0.533	wicked_edge	0.404
atheism	0.529	guitar	0.406
lgbt	0.527	gamedeals	0.406
TumblrInAction	0.524	buildapc	0.407
islam	0.523	techsupport	0.410
SubredditDrama	0.520	travel	0.410

Table 3: Subreddits with the highest and lowest dogmatism scores. Politics and religion are common themes among the most dogmatic subreddits, while hobbies (e.g., photography, homebrewing, buildapc) show the least dogmatism.

We present the results of this analysis in Table 3. The subreddits with the highest levels of dogmatism tend to be oriented around politics and religion (*DebateAChristian* or *ukpolitics*), while those with the lowest levels tend to focus on hobbies (*photography* or *homebrewing*). The subreddit with the highest average dogmatism level, *cringepics*, is a place to make fun of socially awkward messages, often from would-be romantic partners. Dogmatism here tends to take the form of "how could someone be that stupid" and is directed at the subject of the post, as opposed to other members of the community.

Similarly, *SubredditDrama* is a community where people come to talk about fights on the internet or social media. These fights are often then extended in discussion, for example: *"If the best you can come up with is that something you did was legal, it's probably time to own up to being an ass."* The presence of this subreddit in our analysis provides a further sanity check that our model is capturing a robust signal of dogmatism.

5.2 How do dogmatic beliefs cluster? (R2)

Dogmatism is widely considered to be a domain-specific attitude (for example, oriented towards religion or politics) as opposed to a deeper personality trait (Rokeach, 1954). Here we use Reddit as a lens to examine this idea more closely. Are users who are dogmatic about one topic likely to be dogmatic about others? Do clusters of dogmatism exist around particular topics? To find out, we examine the re-

Libertarianism	business	conspiracy	science	Christianity	lgbt
Anarcho_Capitalism	Bitcoin	Republican	Christianity	DebateAChristian	feminisms
Bitcoin	economy	conspiritard	relationship_advice	DebateReligion	Equality
ronpaul	entertainment	ronpaul	worldpolitics	science	SubredditDrama
Conservative	TrueReddit	collapse	MensRights	videos	TwoXChromosomes
Android	socialism	guns	IAmA	news	MensRights
ukpolitics	bestof	worldpolitics	TwoXChromosomes	Libertarianism	offbeat
Equality	philosophy	occupywallstreet	WTF	atheism	fffffffuuuuuuuuuuuuu

Table 4: Clusters of subreddits that share dogmatic users. For example, users who are dogmatic on the *conspiracy* subreddit (a place to discuss conspiracy theories) are also likely to be dogmatic on *guns* or *occupywallstreet*.

lationships between subreddits over which individual users are dogmatic. For example, if many users often post dogmatic comments on both the *politics* and *Christianity* subreddits, but less often on *worldnews*, that would suggest *politics* and *Christianity* are linked per a boost in likelihood of individuals being dogmatic in both.

We sample 1000 Reddit users who posted at least once a year between 2007 and 2015 to construct a corpus of 10 million posts that constitute their entire post history. We then annotate these posts using the classifier and compute the average dogmatism score per subreddit per user. For example, one user might have an average dogmatism level of 0.55 for the *politics* subreddit and 0.45 for the *economics* subreddit. Most users do not post in all subreddits, so we track only subreddits for which a user had posted at least 10 times. Any subreddits with an average dogmatism score higher than 0.50 we consider to be a user's dogmatic subreddits. We then count all pairs of these dogmatic subreddits. For example, 45 users have *politics* and *technology* among their dogmatic subreddits, so we consider *politics* and *technology* as linked 45 times. We compute the mutual information (Church and Hanks, 1990) between these links, which gives us a measure of the subreddits that are most related through dogmatism.

We present the results of this analysis in Table 4, choosing clusters that represent a diverse set of topics. For example, *Libertarianism* is linked through dogmatism to other political communities like *Anarcho_Capitalism*, *ronpaul*, or *ukpolitics*, as well as other topical subreddits like *guns* or *economy*. Similarly, people who are dogmatic in the *business* subreddit also tend to be dogmatic in subreddits for *Bitcoin*, *socialism*, and *technology*. Notably, when we apply the same mutual information analysis to links defined by subreddits posted in by the same user, we

Feature	Direction
total user posts	↑
proportion of posts in most active subreddit	↑
number of subreddits posted in	↓
average number of posts in active articles	↓

Table 5: User behavioral features that are positively and negatively associated with dogmatism. ↑ means the feature is positively predictive with dogmatism, and ↓ means the feature is negatively predictive. For example, the more subreddits a user posts in, the less likely they are to be dogmatic. All features are statistically significant ($p < 0.001$).

see dramatically different results. For example, the subreddits most linked to *science* through user posts are *UpliftingNews*, *photoshopbattles*, and *firstworldanarchist*, and *millionairemakers*.

Finally, we see less obvious connections between subreddits that suggest some people may be dogmatic by nature. For example, among the users who are dogmatic on *politics*, they are also disproportionately dogmatic on unrelated subreddits such as *science* ($p < 0.001$), *technology* ($p < 0.001$), *IAmA* ($p < 0.001$), and *AskReddit* ($p < 0.05$), with p-values computed under a binomial test.

5.3 What user behaviors are predictive of dogmatism? (R3)

We have shown dogmatism is captured by many linguistic features, but can we discover other high-level user behaviors that are similarly predictive?

To find out, we compute metrics of user behavior using the data sample of 1000 users and 10 million posts described in Section 5.2. Specifically, we calculate (1) *activity*: a user's total number of posts, (2) *breadth*: the number of subreddits a user has posted in, (3) *focus*: the proportion of a user's posts that appear in the subreddit where they are most active, and (4) *engagement*: the average number of posts a user contributes to each discussion they engage in.

We then fit these behavioral features to a linear regression model where we predict each user's average dogmatism level. Positive coefficients in this model are positively predictive of dogmatism, while negative coefficients are negatively predictive.

We find this model is significantly predicitive of dogmatism ($R^2 = 0.1$, $p < 0.001$), with all features reaching statistical significance ($p < 0.001$). *Activity* and *focus* are positively associated with dogmatism, while *breadth* and *engagement* are negatively associated (Table 5). Together, these results suggest dogmatic users tend to post frequently and in specific communities, but are not as inclined to continue to engage with a discussion, once it has begun.

5.4 How does dogmatism impact a conversation? (R4)

How does interacting with a dogmatic comment impact a conversation? Are users able to shrug it off? Or do otherwise non-dogmatic users become more dogmatic themselves?

To answer this question, we sample 600,000 conversations triples from Reddit. These conversations consist of two people (A and B) talking, with the structure: A1 → B → A2. This allows us to measure the impact of B's dogmatism on A's response, while also controlling for the dogmatism level initially set by A. Concretely, we model the impact of dogmatism on these conversations through a linear regression. This model takes two features, the dogmatism levels of A1 and B, and predicts the dogmatism response of A2. If B's dogmatism has no effect on A's response, the coefficient that corresponds to B will not be significant in the model. Alternatively, if B's dogmatism does have some effect, it will be captured by the model's coefficient.

We find the coefficient of the B feature in the model is positively associated with dogmatism ($p < 0.001$). In other words, engagement with a dogmatic comment tends to make a user more dogmatic themselves. This effect holds when we run the same model on data subsets consisting only of dogmatic or non-dogmatic users, and also when we conservatively remove all words used by B from A's response (i.e., controlling for quoting effects).

6 Related Work

In contrast to the computational models we have presented, dogmatism is usually measured in psychology through survey scales, in which study participants answer questions designed to reveal underlying personality attributes (Rokeach, 1954). Over time, these surveys have been updated (Shearman and Levine, 2006) and improved to meet standards of psychometric validity (Crowson, 2009).

These surveys are often used to study the relationship between dogmatism and other psychological phenomena. For example, dogmatic people tend to show an increased tendency for confrontation (El-Nawawy and Powers, 2010) or moral conviction and religiosity (Swink, 2011), and less likelihood of cognitive flexibility (Martin et al., 2011), even among stereotypically non-dogmatic groups like atheists (Gurney et al., 2013). From a behavioral standpoint, dogmatic people solve problems differently, spending less time framing a problem and expressing more certainty in their solution (Lohman, 2010). Here we similarly examine how user behaviors on Reddit relate to a language model of dogmatism.

Ertel sought to capture dogmatism linguistically, though a small lexicon of words that correspond with high-level concepts like certainty and compromise (1985). McKenny then used this dictionary to relate dogmatism to argument quality in student essays (2005). Our work expands on this approach, applying supervised models based on a broader set of linguistic categories to identify dogmatism in text.

Other researchers have studied topics similar to dogmatism, such as signals of cognitive style in right-wing political thought (Van Hiel et al., 2010), the language used by trolls on social media (Cheng et al., 2015), or what makes for impartial language on twitter (Zafar et al., 2016). A similar flavor of work has examined linguistic models that capture politeness (Danescu-Niculescu-Mizil et al., 2013), deception (Ott et al., 2011), and authority (Gilbert, 2012). We took inspiration from these models when constructing the feature sets in our work.

Finally, while we examine what makes an opinion dogmatic, other work has pushed further into the structure of arguments, for example classifying their justifications (Hasan and Ng, 2014), or what makes an argument likely to win (Tan et al., 2016). Our

model may allow future researchers to probe these questions more deeply.

7 Conclusion

We have constructed the first corpus of social media posts annotated with dogmatism scores, allowing us to explore linguistic features of dogmatism and build a predictive model that analyzes new content. We apply this model to Reddit, where we discover behavioral predictors of dogmatism and topical patterns in the comments of dogmatic users.

Could we use this computational model to help users shed their dogmatic beliefs? Looking forward, our work makes possible new avenues for encouraging pro-social behavior in online communities.

References

Justin Cheng, Cristian Danescu-Niculescu-Mizil, and Jure Leskovec. 2015. Antisocial behavior in online discussion communities. *arXiv preprint arXiv:1504.00680*.

Kenneth Ward Church and Patrick Hanks. 1990. Word association norms, mutual information, and lexicography. *Computational linguistics*, 16(1):22–29.

H Michael Crowson. 2009. Does the dog scale measure dogmatism? another look at construct validity. *The Journal of social psychology*, 149(3):365–383.

Cristian Danescu-Niculescu-Mizil, Moritz Sudhof, Dan Jurafsky, Jure Leskovec, and Christopher Potts. 2013. A computational approach to politeness with application to social factors. *arXiv preprint arXiv:1306.6078*.

Shayan Doroudi, Ece Kamar, Emma Brunskill, and Eric Horvitz. 2016. Toward a learning science for complex crowdsourcing tasks. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 2623–2634. ACM.

Mohammed El-Nawawy and Shawn Powers. 2010. Al-jazeera english a conciliatory medium in a conflict-driven environment? *Global Media and Communication*, 6(1):61–84.

S Ertel. 1985. Content analysis: An alternative approach to open and closed minds. *The High School Journal*, 68(4):229–240.

Eric Gilbert. 2012. Phrases that signal workplace hierarchy. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, pages 1037–1046. ACM.

Daniel J Gurney, Shelley McKeown, Jamie Churchyard, and Neil Howlett. 2013. Believe it or not: Exploring the relationship between dogmatism and openness within non-religious samples. *Personality and Individual Differences*, 55(8):936–940.

Kazi Saidul Hasan and Vincent Ng. 2014. Why are you taking this stance? identifying and classifying reasons in ideological debates. In *EMNLP*, pages 751–762.

Andrew F Hayes and Klaus Krippendorff. 2007. Answering the call for a standard reliability measure for coding data. *Communication methods and measures*, 1(1):77–89.

Margaret C Lohman. 2010. An unexamined triumvirate: dogmatism, problem solving, and hrd. *Human Resource Development Review*.

Matthew M Martin, Sydney M Staggers, and Carolyn M Anderson. 2011. The relationships between cognitive flexibility with dogmatism, intellectual flexibility, preference for consistency, and self-compassion. *Communication Research Reports*, 28(3):275–280.

Michael McCluskey and Jay Hmielowski. 2012. Opinion expression during social conflict: Comparing online reader comments and letters to the editor. *Journalism*, 13(3):303–319.

John McKenny. 2005. Content analysis of dogmatism compared with corpus analysis of epistemic stance in student essays. *Information Design Journal & Document Design*, 13(1).

Myle Ott, Yejin Choi, Claire Cardie, and Jeffrey T Hancock. 2011. Finding deceptive opinion spam by any stretch of the imagination. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*, pages 309–319. Association for Computational Linguistics.

English Oxford Dictionary. 2016. Definition of dogmatism.

James W Pennebaker and Martha E Francis. 1996. Cognitive, emotional, and language processes in disclosure. *Cognition & Emotion*, 10(6):601–626.

James W Pennebaker, Martha E Francis, and Roger J Booth. 2001. Linguistic inquiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, 71:2001.

James W Pennebaker. 2011. The secret life of pronouns. *New Scientist*, 211(2828):42–45.

Milton Rokeach. 1954. The nature and meaning of dogmatism.

Sachiyo M Shearman and Timothy R Levine. 2006. Dogmatism updated: A scale revision and validation. *Communication Quarterly*, 54(3):275–291.

Victor S Sheng, Foster Provost, and Panagiotis G Ipeirotis. 2008. Get another label? improving data quality and data mining using multiple, noisy labelers. *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 614–622.

Nathan Swink. 2011. Dogmatism and moral conviction in individuals: Injustice for all.

Chenhao Tan, Vlad Niculae, Cristian Danescu-Niculescu-Mizil, and Lillian Lee. 2016. Winning arguments: Interaction dynamics and persuasion strategies in good-faith online discussions. In *Proceedings of WWW*.

Alain Van Hiel, Emma Onraet, and Sarah De Pauw. 2010. The relationship between social-cultural attitudes and behavioral measures of cognitive style: A meta-analytic integration of studies. *Journal of personality*, 78(6):1765–1800.

Markus Wolf, Jan Sedway, Cynthia M Bulik, and Hans Kordy. 2007. Linguistic analyses of natural written language: Unobtrusive assessment of cognitive style in eating disorders. *International Journal of Eating Disorders*, 40(8):711–717.

Kyung-Hyan Yoo and Ulrike Gretzel. 2009. Comparison of deceptive and truthful travel reviews. *Information and communication technologies in tourism 2009*, pages 37–47.

Muhammad Bilal Zafar, Krishna P Gummadi, and Cristian Danescu-Niculescu-Mizil. 2016. Message impartiality in social media discussions. In *Tenth International AAAI Conference on Web and Social Media*.