Proceedings of the Twelfth International AAAI Conference on Web and Social Media (ICWSM 2018)

Conspiracies Online: User Discussions in a Conspiracy Community Following Dramatic Events

Mattia Samory, Tanushree Mitra
Department of Computer Science, Virginia Tech
{samory,tmitra}@vt.edu

Abstract

Online communities play a crucial role in disseminating conspiracy theories. New theories often emerge in the aftermath of catastrophic events. Despite evidence of their widespread appeal, surprisingly little is known about who participates in these event-specific conspiratorial discussions or how do these discussions evolve over time. We study r/conspiracy, an active Reddit community of more than 200,000 users dedicated to conspiratorial discussions. By focusing on four tragic events and 10 years of discussions, we find three distinct user cohorts: *joiners*, who never participated in Reddit but joined r/conspiracy only after the event; *converts* who were active Reddit users but joined r/conspiracy only after the event; and *veterans*, who are longstanding r/conspiracy members. While *joiners* and *converts* have a shorter lifespan in the community in comparison to the *veterans*, *joiners* are more active during their shorter tenure, becoming increasingly engaged over time. Finally, to investigate how these events affect users' conspiratorial discussions, we adopted a causal inference approach to analyze user comments around the time of the events. We find that discussions happening after the event exhibit signs of emotional shock, increased language complexity, and simultaneous expressions of certainty and doubtfulness. Our work provides insight on how online communities may detect new conspiracy theories that emerge ensuing dramatic events, and in the process stop them before they spread.

Introduction

Dramatic events, such as terrorist attacks or mass shootings, cause shock and uncertainty in the population at large. The reasons that lead to these events are often intrinsically complex, and early press coverage may lack clear and definitive evidence. It is no surprise that rumors and speculations surrounding these events naturally arise, often as attempts to rationalize the underlying complex phenomena or to deal with associated feelings of powerlessness (Sunstein and Vermeule 2009). Among these speculations, conspiracy theories are especially dangerous. Conspiracy theories are attempts at offering alternative explanations to significant events by referencing false or unsubstantiated claims (Sunstein and Vermeule 2009). What makes them particularly dangerous is their self-sealing quality—any attempt to correct conspiratorial claims can be folded into the theory itself—making them extremely resistant to correction. Although one could dismiss them as a deviant phenomenon, half of the American population believes in one or more conspiracy theories (Oliver and Wood 2014).

Often the inception and diffusion of conspiracy theories happen online, in well-knit communities of users (Bessi et al. 2015). Without vetting these conspiratorial ideas with accurate information, individuals may react—sometimes radically—to them. One example is the recent shooting by a man who took it upon himself to investigate #pizzagate, a fictitious online conspiracy theory that went viral during the 2016 US Presidential election[1]. It is therefore imperative to understand how users in online communities engage in conspiratorial ideation at the onset of extraordinary events. While some work has investigated conspiratorial narratives on specific social media sites, for example Facebook pages discussing conspiracy news (Bessi et al. 2015), most studies either report analysis from a single instance of a catastrophic event (Koutra, Bennett, and Horvitz 2015) or track user discussions after the event (Starbird 2017). To obtain an in-depth understanding of user behavior at the time a new conspiratorial idea gets introduced, we need large-scale longitudinal analysis of user behavior as new theories surrounding an event take shape. This can help develop effective strategies to counter conspiracism.

In this paper, we examine users who discuss dramatic events in a popular online conspiracy theory community—r/conspiracy. We use a causal inference based longitudinal analysis approach to study the evolution of users' tenures in the community and the effects of the events on their discussion dynamics. This allows us to address several questions. First, do dramatic events draw new users to the conspiracy theory community, and are these users different from longstanding members? Does user engagement with the community fade over time, once the recency of an event fades? Finally, how do dramatic events affect user behavior and reactions at their outset? To answer these questions, we rely on over 6 million comments spanning over 10 years of discussions in r/conspiracy—a community of more than 200,000 users dedicated to conspiracy theorizing. We track the inception of new conspiracy theories by focusing on discussions surrounding four dramatic events. We detect users that dis-

[1] http://wapo.st/2mZBNtT

cuss the events on r/conspiracy, analyze their entire activity histories on Reddit, compare and contrast their long-term engagement within the conspiracy theorizing community, and layout changes in user behavior after exposure to the emerging conspiracy theories.

Types of users that discuss dramatic events. We identify three key user cohorts: *joiners*, who join Reddit to discuss the events; *converts* are existing Reddit users that eschewed r/conspiracy before the events; and *veterans* are longstanding members. We find statistically significant differences among cohorts. Veterans are established and prolific Reddit users, especially dedicated to discussion in r/conspiracy, and in few other related communities. Converts, while similar to veterans for their long and abundant contribution histories on Reddit, seem to be only marginally active in r/conspiracy. Finally, joiners, despite the shorter timespan of their activity, craft richer contributions at a higher rate, and become intensely active over time. This suggests that although dramatic events attract users from both within and outside the platform to discuss emerging conspiracy theories, new adoptees from outside have higher likelihood of becoming proficient members of the community.

Long-term engagement in the community. We devise two indices to quantify user engagement in conspiratorial discussion communities, both inside r/conspiracy and outside in other similar conspiracy subreddits. We find that joiners remain highly engaged well after the dramatic events. This gives additional evidence that joiners would eventually become proficient members of the conspiracist community. Furthermore, longitudinal analyses show that although both veterans and converts become more engaged with time, veterans mature into highly engaged members of the community, while converts engage only mildly. This suggests that veterans become increasingly devoted to conspiracy theorizing with increased participation both within r/conspiracy and other related conspiracy communities.

Changes in user discussions following dramatic events. Finally, we employ interrupted time series analysis—a causal inference method drawn from epidemiology—to assess changes in user behavior two months preceding and following the selected events. We find consistent patterns in how user behavior changes in the vicinity of the four events. Although engagement within r/conspiracy significantly increases after the events, discussions exhibit more negative sentiment, lower content quality, and signs of doubt and rationalization. These results offer valuable insights on the processes of conspiracy theory ideation and adoption.

Literature review

We briefly report on two broad areas of research relevant to this study: the analysis of individual factors associated with belief in conspiracy theories, and the study of dramatic events unfolding in social media.

Individual factors in conspiracy theory belief. Scholars studying the psychological factors behind conspiratorial beliefs, suggest anxiety, paranoia, and feelings of powerlessness as the key correlates (see Grzesiak-Feldman 2013). Another line of research situates conspiratorial beliefs in the domain of logic (Clarke 2002; Garrett and Weeks 2017). In particular, they suggest that the modern conspiracist's logical flaws stem from restricted access to information– a phenomenon called "intellectual isolation", where people's prior views are strengthened and alternative viewpoints are infrequently encountered (Koutra, Bennett, and Horvitz 2015). Indeed, conspiracy theories, like rumors and memes, often appeal to users situated in these ideological eco-chambers (Heath, Bell, and Sternberg 2001; Smith and Leiserowitz 2012). Moreover, the emotional shock that ensues following a dramatic event may heighten the precursors of conspiratorial ideation. Drawing upon these studies, we investigate user dynamics and behavioral changes in a popular conspiracy theorizing community soon after the occurrence of four extraordinary events.

Analyzing reactions to dramatic events in social media. Studying the production and propagation of conspiracy narratives online, has attracted the attention of the scholarly community for several years—(Swami et al. 2011) in the past and more recently (Starbird 2017). Studies have also shown that relying on social media as news sources increases the likelihood of conspiracist ideation (Stempel, Hargrove, and Stempel 2007). Despite ample evidence of social media fueling conspiracy narratives at the onset of dramatic events, we still know very little about how these theories form, evolve, and affect users discussing them. Closest to this study is work analyzing longitudinal changes in online discussions following dramatic events (Lin and Margolin 2014; Maddock et al. 2015). Dramatic events correspond to bursts in user activity, and conspiracy theory-related activities seem to follow recognizable patterns. Natural language expressions—that reflect personality and cognitive processes (Tausczik and Pennebaker 2010)—are especially useful to characterize how user behavior changes after the events. Several works investigate changes in public opinion on gun rights in the aftermath of the Sandy Hook shooting (Koutra, Bennett, and Horvitz 2015; Benton et al. 2016). In particular, language reveals how online discussions by anti-vaccination proponents exhibit conspiratorial world-views (Mitra, Counts, and Pennebaker 2016). Findings from this line of research are possibly unique to each event under study, whereas our current study aims to uncover consistent patterns across different events.

Data preparation

Our data preparation phase has multiple steps (see Fig. 1).

Data collection. We study users of r/conspiracy, a subreddit (i.e., Reddit community) which counts over 500K subscribers to date. Although it features discussions on a range of conspiracy theories, it is especially well-suited to studying those on dramatic events. In fact, the subreddit's own

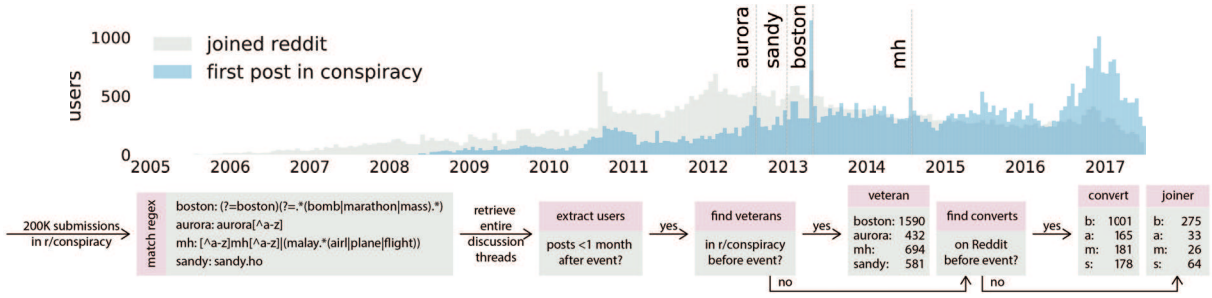

Figure 1: (Top half) Dramatic events correspond to a high number of newcomers in r/conspiracy. Among r/conspiracy users, more users post for the first time in r/conspiracy than on Reddit after the Boston bombing. This implies that there must be a significant number of *converts* in addition to *joiners* who enter the community shortly after the event. (Bottom half) Flowchart detailing how we identify cohorts of users discussing the four events.

definition of conspiracy theory is "a hypothesis that some covert but influential organization is responsible for a circumstance or event"—and indeed a significant fraction of discussions revolve around dramatic events.

We obtained the full set of discussions in r/conspiracy from a publicly available Reddit dump[2]. This dataset spans from the inception of the subreddit in January 2008 to August 2017. The same source also provided us with the entire contribution history of r/conspiracy users on Reddit, consisting of their comments and submissions outside of r/conspiracy. The resulting datasets includes over 200K active users, authoring over 10M submissions and 500M comments overall, of which almost 200K and 6M respectively are in r/conspiracy.

Selecting discussions on dramatic events. Prior work has linked bursts in user activity to discussions on emerging rumors (Maddock et al. 2015; Mitra, Wright, and Gilbert 2017). This inspired us to look at peaks in user activity to detect dramatic events. Figure 1 shows the time when r/conspiracy users post for the first time on Reddit and in r/conspiracy. Both distributions peak at the outset of several dramatic events. After compiling a list of dramatic events matching these peaks, we select four events: the Sandy Hook shooting, the Aurora theater shooting, the take down of Malaysia Airlines flight MH17, and the Boston marathon bombing. We focus on these events for three reasons: 1) r/conspiracy offers high volume of posts at their outset; 2) they are sufficiently distant in time from one another, mitigating the concerns of overlaps; 3) they are semantically diverse: their theorized conspiratorial agents span from private individuals to intergalactic spacecrafts.

Then, we select discussions that refer the four events. We extract submissions in r/conspiracy, posted after each event, and whose title matches the regular expressions in Figure 1. We verify through random sampling that a high fraction of the selected submissions refer to the events. We gather the entire comment threads for the selected submissions, and consider all users appearing in either comments or submissions as discussing the events.

Cohorts of users discussing events. While some events are the subject of conspiracy theories that will be discussed for years to come, such as the 9/11 attacks, we are interested in how newly conceived conspiracy theories develop and impact user behavior. Therefore, we only focus on users that discuss the events, and whose earliest contribution to those discussions dates within one month after the event. Moreover, previous exposure to conspiracy theories, as well as familiarity with community norms, may influence user expressions. Therefore, we divide users into three cohorts: *veterans*, who posted in r/conspiracy before the event; *converts*, existing Reddit users that join r/conspiracy after the event; and *joiners*, who join Reddit one month after the event.

Long-term adopters of conspiracy theories typically believe in multiple, often contradictory theories (Sunstein and Vermeule 2009). Veterans allow us to study this segment of the online conspiracy community because they discuss conspiracy theories before the events, in addition to the ones following the events. Moreover, the need to adapt to community conventions may affect expressions of newcomers (Kiene and Hill 2016). For this reason, we distinguish converts—users who are new to r/conspiracy but who are accustomed to Reddit and thus face relatively low entry barriers—from joiners, novice Reddit users. Figure 1 reports the cardinality of user cohorts across events.

Finding communities similar to r/conspiracy. Given the broad scope of r/conspiracy, users may take discussions about specific conspiracy theories to dedicated communities, such as r/flatearth. To analyze users' activity in these ancillary communities, we first need to assess whether they are related to r/conspiracy. To this end, we adopt a data-driven measure of semantic similarity (Martin 2017). We estimate how much closer a community is to r/conspiracy than to a community for scientific debate—a polar opposite to conspiracy theory discussion (Bessi et al. 2015). We choose the popular r/science as a champion scientific community. We collect the timelines of all users in r/conspiracy and r/science with at least 10 contributions in the two subreddits (Reddit limits scraping to 1000 contributions per user). For each pair

[2]available on Google BigQuery: https://bigquery.cloud.google.com/dataset/fh-bigquery:reddit

of subreddits in the user timelines, we assess their similarity by measuring whether they share a surprisingly high number of users. We discard pairs sharing less than 10 users. Finally, we measure surprise through positive pointwise mutual information. We use this score to measure similarity between a community and r/conspiracy. We inspect the 50 communities that are the most and least similar to r/conspiracy, and verify that the measure follows human judgment. Table 1 reports the 10 most similar subreddits ($sim\,(s, c)$ = similarity between subreddit s and r/conspiracy).

Types of users who discuss dramatic events

The previous section described how we identify users that discuss emerging conspiracy theories, and how they can be divided into cohorts. How do these cohorts differ? In this section, we introduce measures to characterize user activity and compute differences across cohorts.

Measuring user profiles. We extract higher-level features that summarize the activity of users throughout their permanence on Reddit. Prolonged exposure to conspiracy theory discussion may affect users' propensity to embrace new conspiracy theories (Oliver and Wood 2014). A simplistic measure of exposure is user *lifespan* in r/conspiracy—the number of seconds between users' first and last available contribution in r/conspiracy. *Lifespan* measures how long—but not how much—users are active in the community. Therefore, we measure their *contribution volume* in r/conspiracy, as the total number of submissions and comments made by the user. The content of users' discussions also reflects how much effort they put into their contributions. Are they highly verbose and diverse in their contribution? Or are they simply repeating themselves? We measure users' *verbosity* as the average number of characters in user comments. We measure *redundancy* in their contribution by finding gzip compression ratio (Stamatatos 2009) of each user's comment corpus. The more one user contributes comments that are similar to one another, the higher the compression ratio. How does a user's activity in r/conspiracy compare to his overall tenure on Reddit? To put a user's activity in r/conspiracy into perspective of his overall tenure on Reddit, we compute user's *lifespan* and *contribution volume* in all of Reddit. Next, we observe the multiple *memberships* that a user may have by finding the number of subreddits that a user actively contributes to. Are users highly active in select communities such as r/conspiracy while only posting occasionally in other communities? To answer this, we compute *membership exclusivity*—the disproportion in user participation across communities—through the Gini index of subreddit activity.

Who discusses dramatic events in r/conspiracy? Are there characteristic differences based on when user enter the community? Here we compare and contrast the characteristics of our three cohorts across the measures outlined in the previous section. To validate the statistical differences, we employ analysis of variance (ANOVA), followed by post-hoc Tukey's Honest Significant Difference (HSD) tests, to

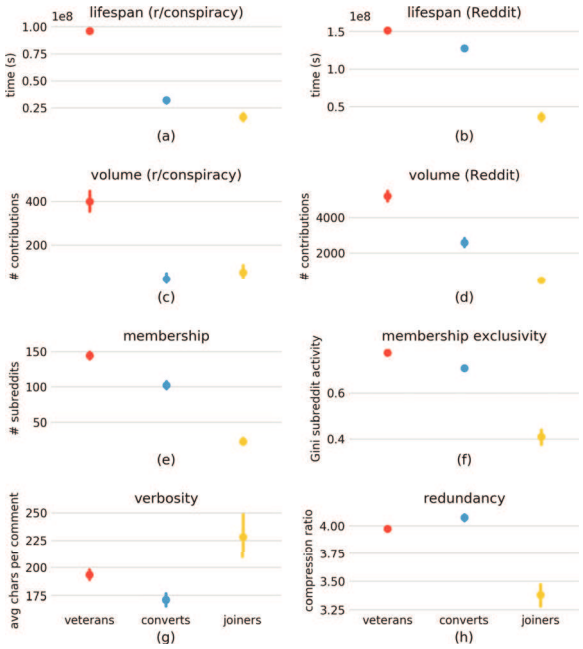

Figure 2: Differences between user cohorts. There is no statistical difference in the contribution volume in r/conspiracy between joiners and converts, despite the shorter lifespan of joiners. Joiners also put the most effort in their contribution, crafting longer, less repetitive comments.

account for multiple comparisons: all results presented, unless otherwise specified, are significant at $p < 0.01$ for the ANOVA test, and $p < 0.05$ for the Tukey test. Figure 2 shows the means and 95% confidence intervals of user features across cohorts.

Joiners contribute the most *verbose* and the least *redundant* comments, followed by veterans and converts in order. Prior work suggests that this verbosity and lexical variety may be due to the numerous interpretations and pieces of evidence that are part of conspiracy theorizing (Maddock et al. 2015). Through this lens, joiners and veterans are the most involved in theorizing discussions.

Activity *lifespan* and *contribution volume* help better understand these differences. Veterans remain active for the longest time on Reddit, and specifically in r/conspiracy (Figure 2b and 2a). The *volume of contribution* also follows similar trend. Conversely, the joiner cohort is the one with the shortest *lifespan* on Reddit, and in r/conspiracy. This may be unsurprising, considering that joiners become active on Reddit later, and in r/conspiracy no sooner than other cohorts. Ultimately, joiners have the least number of contributions on Reddit. However, despite the much longer *lifespan* of converts, there is no statistically significant difference in the *volume of contributions* between joiners and converts.

Finally, we contextualize activity in r/conspiracy with participation in other subreddits. User activity naturally correlates with the number of subreddits that they are active in. In order, joiners, converts, and veterans participate in increasing numbers of subreddits, i.e, *membership* of joiners

Top 10 subreddits most similar to r/conspiracy

CHEMPRINTS	Bilderberg
conspiracyhub	greenlight2
WhiteNationalism	greenlight
HealthConspiracy	OccupyLangley
mysterybabylon	moonhoax

Table 1: Subreddits that share the most surprising number of users with r/conspiracy, sorted in decreasing order of their similarity score ($sim\,(s, c)$)

$<$ converts $<$ veterans (Figure 2e). Surprisingly, this trend also replicates in how unequally users distribute their activity across the subreddits (Pearson's $\rho = 0.42$, $p < 0.001$). In other words, joiners contribute to the fewest subreddits (low in *membership*), but are the most similarly active in all of them (low Gini coefficient denotes low inequality). Veterans despite being active in multiple subreddits (high *membership*), still focus their attention on only a few select communities. The same is true for converts. One possible explanation is that users' interests do not scale with the number of subreddits they follow. Instead, they tend to focus their activity on a limited number of subreddits. Do these select few subreddits also foster conspiratorial discussions? We explore this possibility in the next section.

These results sketch a potential profile of the user cohorts. Veterans are established and prolific Reddit users, especially dedicated to discussions in r/conspiracy, and in few other communities. Converts, while similar to veterans for their long and prolific tenure on Reddit, put lesser effort (lowest *verbosity* and highest *redundancy*), and are only marginally active in r/conspiracy. Finally, joiners, despite the shorter timespan at their disposal (lowest *lifespan*), craft richer contributions (highest *verbosity* and lowest *redundancy*). Building on these observations, next, we discuss how engagement in the conspiracy theory community evolves over time.

Long-term engagement with the community

Does user's engagement with the conspiracy theorizing community change over time following a dramatic event? To answer this question, we first devise measures to quantify user's engagement and then map their evolution over the user's lifespan.

Measuring engagement. Do users who follow a popular conspiracy community (r/conspiracy) predominantly contribute to this single community? To investigate this question, we devise our first index, *engagement$_{in}$* as a rank normalized measure of user's contribution in r/conspiracy. Higher ranks correspond to more contributions. More formally, for each user u:

$$engagement^{u}_{in} = \frac{rank\,(a^{u}_{c})}{|S^{u}|} \tag{1}$$

where S^u is the set of user u's subreddits, a^u_i is her contributions to subreddit i, and c is r/conspiracy.

While r/conspiracy is the largest conspiracy discussion community in Reddit, the site hosts multiple other subreddits for discussing specific conspiracy theories. Thus, it is important to capture user's engagement with conspiratorial discussion communities beyond r/conspiracy. Therefore, we devise our second index, *engagement$_{out}$*, which measures how much of users' activities outside r/conspiracy are taking place in similar conspiratorial communities. First, we compile a list of communities related to r/conspiracy based on their shared user base. We detail our similarity score calculation method in the earlier data preparation section (see Table 1 for a list of top scored subreddits). Next, we compute the fraction of the user's activity in each community, weighted by the community's similarity to r/conspiracy:

$$engagement^{u}_{out} = \frac{\sum_{s \in S^u \setminus \{c\}} (sim\,(s, c) \cdot a^{u}_{s})}{\sum_{s \in S^u \setminus \{c\}} a^{u}_{s}} \tag{2}$$

where $sim\,(s, c)$ is the similarity between subreddit s and r/conspiracy. Quite surprisingly, we find that the two engagement measures correlate (Pearson's $\rho = 0.14$, $p < 0.001$). Recall that users with high *engagement$_{in}$* already have a large fraction of their contribution in r/conspiracy. The positive correlation between the two engagement measures implies that their remaining activity is also in similar subreddits, suggesting that r/conspiracy users tend to contribute to similar conspiratorial communities. In particular, veterans show higher correlation ($\rho = 0.19$, $p < 0.001$), while the correlation is lower for converts ($\rho = 0.07$, $p < 0.01$), and is not statistically significant for joiners ($\rho = 0.10$, $p > 0.05$).

How does engagement evolve after the events? We investigate how cohorts engage with the conspiracist community within and outside of r/conspiracy, throughout their lifespan. For each user, we divide their comments in time segments corresponding to 10% of their lifespan, and compute the mean user engagement in each decile. Figure 3 shows the evolution of *engagement$_{in}$* and *engagement$_{out}$* throughout users' lifespans.

First, we look at how engaged users will ultimately be, after their entire lifespan. The figure shows these values at its rightmost end. Joiners are highly engaged in the conspiracist community both inside and outside r/conspiracy, although veterans sport the highest values of engagement in r/conspiracy specifically. Converts are the least engaged both within and outside r/conspiracy. The temporal evolution of *engagement$_{in}$* further illustrates the differences between joiners and veterans within r/conspiracy. Joiners remain highly engaged throughout their lifespan. Veterans, on the other hand, show relatively low engagement in the beginning, and become more engaged with time. Converts follow a similar trajectory, but remain less engaged throughout their lifespan. Finally, engagement in related communities (*engagement$_{out}$*) increases with time for all three cohorts, wherein joiners show higher values than veterans, and veterans of converts.

It seems joiners are "born," veterans are "made," converts "never become" highly engaged members of the conspiracy discussion community. This reinforces the results

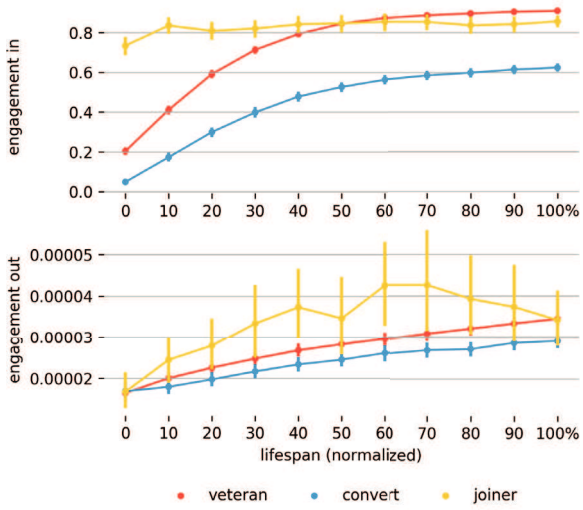

Figure 3: How $engagement_{in}$ and $engagement_{out}$ evolve during the lifespan of users in different cohorts. Veterans and joiners reach high engagement levels—with the difference that joiners are highly engaged from the outset. Although converts also become more engaged with time, they remain less engaged than veterans and joiners. $engagement_{out}$ increases for all three cohorts.

in the previous section that suggested joiners and veterans put significant effort in their contributions. Overall, we find increase in engagement throughout users' lifespans—which is consistent with sociological interpretations of conspiracy theory diffusion (Sunstein and Vermeule 2009). One of the causes may be self-selection. In particular, joiners, who become members of the conspiracy theory discussion community in the wake of dramatic events, may be particularly predisposed to adopt conspiratorial attitudes. Scholars have defined predisposition as a state which when activated makes an individual respond favorably to the stimulus (Rokeach 1968). Another interpretation is increasing susceptibility to radicalization. When individuals are ingrained in these communities over long periods of time, they get exposed to new conspiracy theories. Repeated exposure coupled with limited access to diverse opinions reinforces their existing conspiratorial belief and provides the ideal environment to formulate increasingly radical ideas (Bessi et al. 2015). This may explain why veterans, who are already members of the community, become more engaged as dramatic events generate new conspiracy theories.

Changes in user discussions following dramatic events

Assessing change after dramatic events. To measure change in user behavior following a dramatic event, we employ Interrupted Time Series analysis (ITS in short). ITS is a causal inference approach, widely used in epidemiology and community intervention research to measure the change in user behavior following an experimental intervention. One advantage of ITS over alternative methods is that its quasi-

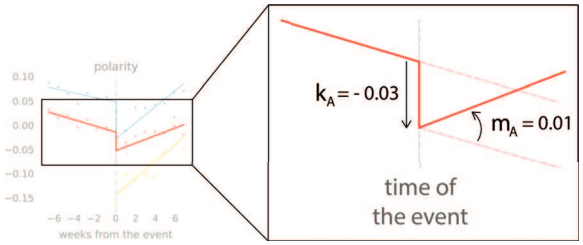

Figure 4: Example interrupted time series model, reporting change in comment polarity for veterans. The model shows a negative shift in polarity at the time of the event ($k_A < 0$), but a positive change in slope ($m_A > 0$) after the event.

experimental design can control for trends in the variable of interest that pre-dated the event, allowing causal interpretation (Biglan, Ary, and Wagenaar 2000). In its simplest form, an ITS model uses linear regression to estimate three parameters: the slope before the intervention, the change in level at the time of intervention, and the change in slope after the intervention (see Figure 4). More formally, to measure the effect of the event on a variable of interest y, our baseline model uses OLS regression to fit the following model:

$$y \sim \underbrace{k + mt}_{\text{before the event}} + \underbrace{\mathcal{H}_T(t)}_{\text{at the event}} \cdot \underbrace{(k_A + m_A t)}_{\text{change after the event}}$$

where t is the time step in weeks from the event, k and m are the coefficients for intercept and slope of the variable before the event, and k_A and m_A are the *change* in level and slope after the event (note that we denote coefficients estimating change after the event with subscript A). $\mathcal{H}_T$ is the step function centered at the time of the event, T.

We also augment the baseline model to control for user cohorts (Simonton 1977). We introduce a variable δ that selects model coefficients corresponding to the user cohorts. We use veterans as the reference class in dummy-coding δ. Using a reference class is essential to avoid multicollinearity and to ensure that the model measures cohort specific effects. By definition, joiners are only active after the event. Hence, their model only includes variables after the event. Note that, in our notation, apexes v, c, and j denote user cohort, respectively veterans, converts, and joiners. The complete model is:

$$\begin{aligned} y \sim k^v + m^v t + \quad & \mathcal{H}_T(t) \cdot (k_A^v + m_A^v t) + \quad \Leftarrow \text{veterans} \\ +\delta^c \cdot (k^c + m^c t + \quad & \mathcal{H}_T(t) \cdot (k_A^c + m_A^c t)) + \quad \Leftarrow \text{converts} \\ +\delta^j \cdot (\quad & \mathcal{H}_T(t) \cdot (k_A^j + m_A^j t)) \quad \Leftarrow \text{joiners} \end{aligned}$$

An alternative to formulating a single model would have been estimating effects on the three cohorts separately. Such a model, however, would not allow cohort specific comparisons. A comprehensive model provides additional control, which allows for quantifying changes that are specific to joiners and converts while controlling for veterans.

Features describing user comments

We employ four sets of features to track changes in user commenting behavior in the conspiratorial communities soon after the occurrence of a dramatic event (see Table 2).

engagement	interaction quality
$\text{engagement}_{\text{in}}$	score
$\text{engagement}_{\text{out}}$	readability
	delay
argumentation	function words
interrogative beginnings	**emotional response**
insight	sentiment polarity
causation	death
certainty	biological processes
assent	inhibition
cognitive mechanisms	I pronoun

Table 2: Measures of comments that we track to understand how users react to events.

Engagement. Does discussing emerging conspiracy theories increase participation in the conspiracy theory community? To answer this question, we trace the evolution of $engagement_{in}$ and $engagement_{out}$ in a time frame surrounding the events. These engagement measures characterize the relationship between users and the conspiracy theory discussion community. In particular, $engagement_{in}$ measures how much more are users active in r/conspiracy, with respect to other communities. $engagement_{out}$, on the other hand, computes how much of the users' activity belongs to other communities that share members with r/conspiracy.

Interaction quality. To assess changes in quality of user comments, we compute four measures. First, we compute *delay* between two consecutive comments made by the same user. *Delay* measures the amount of time users dedicate to crafting contributions. Second, we use SMOG grade to measure *readability* of the content. SMOG estimates the difficulty of comprehending comments and correlates well with human judgments of text clarity—high SMOG implies more complex (Diakopoulos 2015). Next, we use the fraction of *function words* to assess whether comments are heavy on linguistic style or on linguistic content. While *function words* (such as, pronouns, prepositions, articles, conjunctions) denote linguistic style, "content words" (such as nouns and regular verbs) indicate communication content (Tausczik and Pennebaker 2010). Finally, we gauge community feedback through the Reddit *score* of the comments. *Score* captures the difference in upvotes and downvotes received by a comment.

Emotional response. Emotional response features measure the emotions expressed in the user reactions to the events. First, we use VADER (Hutto and Gilbert 2014) to analyze sentiment *polarity*. In addition to sentiment *polarity*, we analyze psycholinguistic measures that offer a more nuanced picture of the emotional expressions in user comments. Next, we resort to the Linguistic Inquiry and Word Count (LIWC) dictionary to analyze psycholinguistic measures in users' reactions to the dramatic events (Tausczik and Pennebaker 2010). Previous work found that in the wake of shooting events, individuals use more expressions of physical harm and powerlessness, and fewer references to one's self with respect to others (Saha and De Choudhury 2017). Therefore, we capture expressions of physical pain using the *death* category, that includes words like "kill", "grief", and "suicide", and the *biological processes* category, that contains words like "blood", "pain", and "doctor". We also compute expressions of powerlessness using the *inhibition* category, that contains words like "block", "constrain", and "stop". Furthermore, we assess self-centeredness—how much users pose attention to themselves with respect to others—with LIWC's self-referential category.

Argumentation. Finally, argumentation features measure how users reason about the events. To this end, we detect the cognitive and dialectic constructs in user comments. We resort to LIWC's *cognitive mechanisms* and its subcategories: *causation*, *certainty*, *insight*, and *assent*. These categories have been extensively used for exploring thought processes (Tausczik and Pennebaker 2010). Furthermore, we detect sentences with *interrogative beginnings*, which potentially question the emerging conspiracy theories.

How dramatic events affect user behavior

Having compiled the list of features representing user's reaction to an event, we now assess the effect of dramatic events on user behavior. We employ ITS analysis (described earlier). We compute the four feature sets for all user comments on Reddit, within a four-month time window surrounding each event. For space constraints, we only display results for the Boston marathon bombing event (see Fig. 5).

How does engagement within the conspiracist community change? By using $engagement_{in}$ and $engagement_{out}$ respectively, we assess how engagement in r/conspiracy, and in other related communities, change in the wake of the events. Veterans, in particular, show an increase in $engagement_{in}$ at the time of the event, and a small, negative change in slope in the following weeks. By definition, joiners and converts do not have values for $engagement_{in}$ before the event. We also find that $engagement_{out}$ shows contrasting results. For example, while the $engagement_{out}$ of converts decreases after the Sandy Hook event, it increases after the Boston bombing. One possible explanation is that users may take discussions on an emerging conspiracy theory to a dedicated community, such as r/findbostonbombers, leading to higher $engagement_{out}$. Thus user engagement outside r/conspiracy may vary depending on the availability of an appropriate alternative community. Interestingly, however, $engagement_{out}$ peaks for veterans at the time of the event. Their increased engagement outside the main conspiracy theory community may be an effort to spread the emerging theories, and recruit new users. In particular, future work should verify whether converts had come in contact with veterans outside of r/conspiracy before joining the community.

Formulating conspiracy theories that try to explain dramatic events can be seen as a collective attempt to define what threatens the community. Scholars suggest that taking part in this process should increase the users' sense of belonging, and thus their involvement in the community

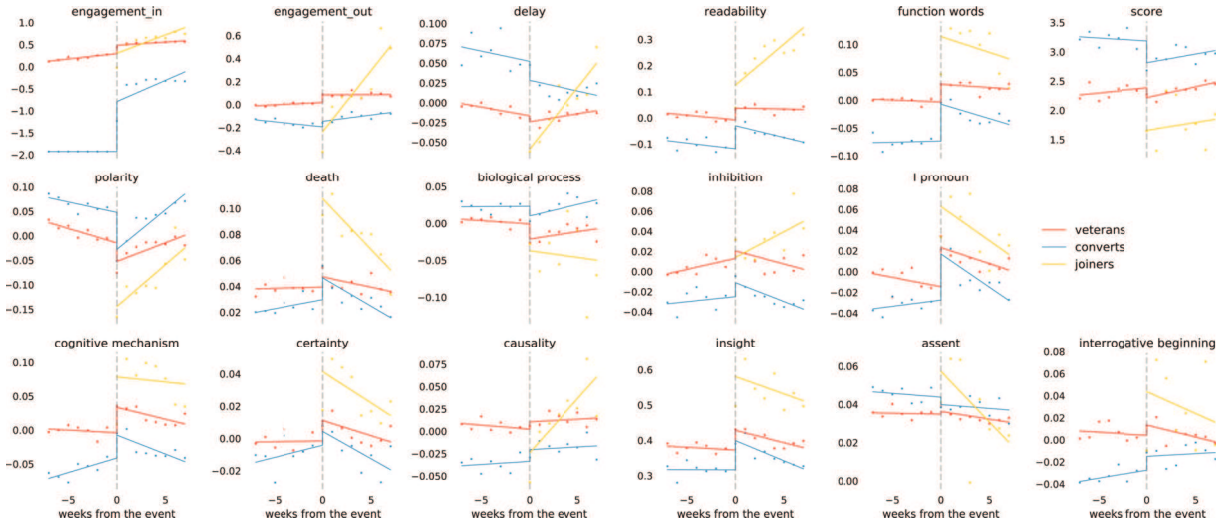

Figure 5: Changes in user behavior after the Boston bombing. Users show increased engagement within the conspiracy community (*engagement*$_{in}$, *engagement*$_{out}$). Although they comment more frequently after the event (*delay*), the quality of their comments worsens: their comments are more complex to read (*readability*, *function words*), and receive more negative feedback (*score*). This may be due to emotional shock. Their increasingly self-referential posts (*I*) express more negative sentiment (*polarity*), pain (*death*, *biological processes*), and powerlessness (*inhibition*). As users evaluate emerging conspiracy theories, their language is simultaneously confirmatory (*certainty*, *causality*, *insight*, *assent*) and doubtful (*interrogative beginning*).

(Goode and Ben-Yehuda 1994). Indeed, we find these manifestations of active participation in the comments with highest *engagement*$_{in}$ at the time of the events:

> *"you are in a good position to spread this info to young people, who are being indoctrinated without knowing it. young people are our future!"*

How does interaction quality change? Users post more hastily after the events. Veterans wait less time between consecutive comments at the time of the event (not statistically significant for the Boston bombing event). Commenting *delay* is also at a local minimum for joiners, and increases with time after the event. Contrary to joiners, converts show a negative change in slope. Moreover, comments become more difficult to read at the time of the event for converts and veterans (except Malaysia Airlines veterans). Text also shows more *function words* at the time of the event for veterans and converts, which further decrease for veterans (not statistically significant). Finally, we analyze the comments' *score*, Reddit's community feedback on comment quality. We find that all cohorts receive more negative feedback at the time of the event.

Overall it seems that while engagement within the conspiracy theory discussion community increases, the quality of comments decreases after the events. It is possible that because of the excitement caused by the dramatic events, users redact comments more hastily, resulting in less clear and less meaningful contributions that are unpopular with the community. Below is an example of an ill-formed, accusatory comment posted in r/conspiracy after the Sandyhook shooting event:

> *"When we call you a paid shill its because we dont believe you can be this stupid on your own so you should take it as a complement"*

Another possibility is that the text is less readable because of higher complexity of content. This interpretation suggests that users express more convoluted arguments as emerging conspiracy theories take shape.

How do expressed emotions change? Both veterans and converts express more negative sentiment at the time of the event, and all cohorts express more positive sentiment as time passes (with the exception of Sandy Hook and Malaysia Airlines joiners). *Death*-related words increase at the time of the event and subside with time for all cohorts (with the exception of Sandy Hook converts). For converts and veterans, *inhibition* words peak at the time of the event and then subside. Conversely, joiners use increasingly more *inhibition* words after all events. Veterans and converts use less *biological processes*, while veterans and joiners show more self-centeredness at the time of the event.

Heightened emotionality and personal involvement may drive individuals to accept conspiracy theories. Conversely, popular conspiracy theories often appeal to emotional rather than rational responses (Clarke 2002; Sunstein 2016). We find that r/conspiracy users show emotional shock after the events. The following comment exemplifies the negativity that follows the events:

> *"[...] I'm saddened like I said. Am I shocked that this s[**]t goes on in a society that almost revolves around and glorifies violence? Not really, sadly. Not really. [...]"*

Especially, cohorts exhibit less interpersonal focus, i.e., they mention themselves more than others in the discussions following the events. This may be because they identify themselves as potential victims of the alleged conspiracy

(Sunstein and Vermeule 2009).

How does argumentation change? Emerging conspiracy theories that aim to explain dramatic events lack consensus at first. We track linguistic markers of argumentation to better understand how users reason about the events and the ensuing conspiracy theories.

Cognitive processes peak at the time of the event, and subside with time (except for Sandy Hook converts). Interestingly, *causation*, *certainty*, *insight*, and sentences with *interrogative beginnings*, all show increases at the time of the event, and subside with time after the event, with few exceptions (e.g., *causality* does not shows positive trends in all Boston cohorts). Furthermore, converts show decreased use of *assent* words (such as, "agree", "OK", "yes") at the time of the event, while joiners' usage of *assent* words drops steadily after the event.

These results corroborate previous scholarly claims that evaluating dramatic events may be cognitively demanding (Swami et al. 2014). Additionally, processing the new conspiracy theories that emerge with them may require arbitration and argumentation as demonstrated by the usage dynamics of *assent*, *certainty* and *interrogative beginnings*. It appears that all cohorts undergo a phase in which they express certainty, but at the same time they introduce more questions in the discourse:

> *"How would you hear about it? I'm sure they would try to make it as secret/non-threatening as possible. Even the nazi's kept the abduction of jews quiet for a little while."*

This apparent contradiction is in line with prior work on conspiracy theorizing, which proposes that the exposure to new and potentially unsettling information strengthens irrational beliefs. Intuitively, this phase of elaboration on the events corresponds to the process of conspiracy theory formation.

Discussion

Dramatic events reinforce the conspiracy theory community. By using a casual inference based approach, we are able to observe the effects that dramatic events have on a community dedicated to conspiracy theory discussions. The increasing trends in user engagement following these events, coupled with r/conspiracy's steady influx of new users and overall rise in user engagement throughout users' lifespans, together suggest that dramatic events have a reinforcing effect on a conspiracy discussion community. Social psychologists have offered two key predispositions that lead people towards conspiratorial ideation. Both are in line with our observations. The first is a propensity to attribute the source of unexplained or extraordinary events to unseen, intentional forces; in other words, a tendency to draw connections between seemingly related phenomena (Oliver and Wood 2014). The events in our study are extraordinary events which are lacking clarity at the time of occurrence. This in turn indicates their proclivity to spur conspiratorial discussions. The second predisposition is a natural attraction toward melodramatic narratives as explanations for prominent events. Scholars have argued that this conspiratorial predisposition is the reason behind why ordinary people adopt conspiracy theories. We do observe heightened anxiety and negative emotion through their persistent usage of *death* related words and expressions of negative sentiment. These negative expressions further amplify at the time of the event for all three user cohorts. These results suggest that dramatic events may trigger users with existing conspiratorial predispositions to embrace new conspiracy theories. Moreover, the steady increase in engagement of new adoptees (i.e. joiners and converts), indicates that dramatic events can nudge new adoptees into becoming well integrated members of the conspiracy community (Rokeach 1968).

Countering conspiracism. Our results suggest that dramatic events may be a particularly fertile ground for the growth of conspiracy theory communities. Thus, institutions aiming to counter conspiracism should focus their efforts on these communities, especially at the time when an ambiguous public event occurs. Our findings lead us to think that these efforts should also target user cohorts differently. We assign users to cohorts in the context of each dramatic event and report significantly different behavioral patterns. Converts, for example, despite being active Reddit users, abstain from r/conspiracy discussions until after the events. Moreover, even after joining the community, their argumentation is the least certain. Thus among all the other cohorts, converts might be the most skeptical of conspiracist explanations. They also show the lowest contribution effort and lowest overall engagement within the conspiracist community. These results imply that in the long term, converts will not become core members of the community. In contrast, joiners are particularly eager r/conspiracy contributors since they join the community despite being new to Reddit and remain engaged throughout their lifespan. Hence, it is likely that joiners will eventually become stable veterans. Thus, organizations working towards dispelling conspiratorial beliefs should focus their efforts on *joiners* over other cohorts. Timely intervention is important because joiners show the most extreme signs of distress at the time of the events and the most radical changes as time after the events passes. The initial phase of elaboration of the emerging theories may be when an intervention is most effective.

Finally, longstanding conspiracy theorist are the most difficult to dissuade—our *veterans*. Previous work suggests cognitive infiltration (disrupting consensus among conspiracists) as the most effective way to counter conspiracism among radicalized believers (Sunstein and Vermeule 2009). We find that veterans show signs of radicalization through their increased—and increasingly exclusive—engagement within the conspiracy theory community. Moreover, discussing emerging conspiracy theories in the wake of dramatic events further boosts their engagement. Therefore, early identification of veterans is crucial for effectively countering conspiratorial ideation. After all veterans only become fully engaged within the conspiracy community in their later years.

Conclusions

This work analyzes users discussing four dramatic events in r/conspiracy, a popular Reddit community dedicated to dis-

cussing conspiracy theories. We find that the time of occurrence of dramatic events corresponds to peak recruitments in the community. Among these recruits, joiners become proficient members of r/conspiracy—more so than converts. We devise measures of engagement with the conspiracy theory community on Reddit at large, and find that joiners and veterans become increasingly engaged with the community throughout their tenure. Then, we quantify longitudinal changes in user behavior in the wake of the events, and find that, although users become more engaged with the community after the events, they contribute comments of lower quality, express emotional distress, and exhibit confused argumentation. Nonetheless, this work has several limitations. Although we apply causal inference methods to assess the effect of events on user behavior, causality claims should be cautiously considered. Moreover, network analysis of the online conspiracist community could also help disentangle the informational and social roles of user cohorts. Finally, this work finds that multiple communities of Reddit relate to conspiracism. An actionable definition of "conspiracist community" would grant better understanding of the movement, and better strategies to counter it.

References

Benton, A.; Hancock, B.; Coppersmith, G.; Ayers, J. W.; and Dredze, M. 2016. After Sandy Hook Elementary: A Year in the Gun Control Debate on Twitter. *Data for Good Exchange*.

Bessi, A.; Coletto, M.; Davidescu, G. A.; Scala, A.; Caldarelli, G.; and Quattrociocchi, W. 2015. Science vs Conspiracy: Collective Narratives in the Age of Misinformation. *PLOS ONE*.

Biglan, A.; Ary, D.; and Wagenaar, A. C. 2000. The Value of Interrupted Time-Series Experiments for Community Intervention Research. *Prev Sci.*

Clarke, S. 2002. Conspiracy Theories and Conspiracy Theorizing. *Philos. Soc. Sci.*

Diakopoulos, N. 2015. Picking the NYT Picks: Editorial Criteria and Automation in the Curation of Online News Comments. *#ISOJ*.

Garrett, R. K., and Weeks, B. E. 2017. Epistemic beliefs' role in promoting misperceptions and conspiracist ideation. *PLOS ONE*.

Goode, E., and Ben-Yehuda, N. 1994. Moral Panics: Culture, Politics, and Social Construction. *Annu. Rev. Sociol*.

Grzesiak-Feldman, M. 2013. The Effect of High-Anxiety Situations on Conspiracy Thinking. *Curr Psychol*.

Heath, C.; Bell, C.; and Sternberg, E. 2001. Emotional selection in memes: The case of urban legends. *J. Pers. Soc. Psychol.*

Hutto, C. J., and Gilbert, E. 2014. VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. In *ICWSM*.

Kiene, C., and Hill, B. M. 2016. Surviving an "Eternal September" How an Online Community Managed a Surge of Newcomers. *CHI*.

Koutra, D.; Bennett, P.; and Horvitz, E. 2015. Events and Controversies: Influences of a Shocking News Event on Information Seeking. *TAIA workshop in SIGIR*.

Lin, Y. R., and Margolin, D. 2014. The ripple of fear, sympathy and solidarity during the Boston bombings. *EPJ Data Science*.

Maddock, J.; Starbird, K.; Al-Hassani, H. J.; Sandoval, D. E.; Orand, M.; and Mason, R. M. 2015. Characterizing Online Rumoring Behavior Using Multi-Dimensional Signatures. In *CSCW*.

Martin, T. 2017. community2vec : Vector representations of online communities encode semantic relationships. In *Workshop on NLP and CSS*.

Mitra, T.; Counts, S.; and Pennebaker, J. W. 2016. Understanding Anti-Vaccination Attitudes in Social Media. In *ICWSM*.

Mitra, T.; Wright, G.; and Gilbert, E. 2017. Credibility and the Dynamics of Collective Attention. *PACM on HCI* 1(CSCW).

Oliver, J. E., and Wood, T. J. 2014. Conspiracy theories and the paranoid style(s) of mass opinion. *Am. J. Political Sci*.

Rokeach, M. 1968. Beliefs, attitudes and values: A theory of organization and change.

Saha, K., and De Choudhury, M. 2017. Modeling Stress with Social Media Around Incidents of Gun Violence on College Campuses. *PACM on HCI* (CSCW).

Simonton, D. K. 1977. Cross-sectional time-series experiments: Some suggested statistical analyses. *Psychol. Bull*.

Smith, N., and Leiserowitz, A. 2012. The Rise of Global Warming Skepticism: Exploring Affective Image Associations in the United States Over Time. *Risk Analysis*.

Stamatatos, E. 2009. A survey of modern authorship attribution methods. *J. Assoc. Inf. Sci. Technol*.

Starbird, K. 2017. Examining the Alternative Media Ecosystem through the Production of Alternative Narratives of Mass Shooting Events on Twitter. In *ICWSM*.

Stempel, C.; Hargrove, T.; and Stempel, G. H. 2007. Media Use, Social Structure, and Belief in 9/11 Conspiracy Theories. *J. Mass Commun. Q*.

Sunstein, C. R., and Vermeule, A. 2009. Conspiracy Theories: Causes and Cures. *J. Political Philos*.

Sunstein, C. R. 2016. *Conspiracy Theories and Other Dangerous Ideas*. Simon and Schuster.

Swami, V.; Coles, R.; Stieger, S.; Pietschnig, J.; Furnham, A.; Rehim, S.; and Voracek, M. 2011. Conspiracist ideation in Britain and Austria: Evidence of a monological belief system and associations between individual psychological differences and real-world and fictitious conspiracy theories. *Br. J. Psychol.*

Swami, V.; Voracek, M.; Stieger, S.; Tran, U. S.; and Furnham, A. 2014. Analytic thinking reduces belief in conspiracy theories. *Cognition*.

Tausczik, Y. R., and Pennebaker, J. W. 2010. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *J. Lang. Soc. Psychol*.