

This accepted version of the article may differ from the final published version.  
This is an Accepted Manuscript for *Disaster Medicine and Public Health Preparedness* as  
part of the Cambridge Coronavirus Collection  
DOI: 10.1017/dmp.2020.404

The Structure of Tweets about Vaccine Safety Between Health Organizations, Experts  
and the Public: Analyzing Risk Communication Conversations

David Manheim<sup>1,\*</sup>, Anat Gesser-Edelsburg<sup>1</sup>

<sup>1</sup>Health and Risk Communication Research Center University of Haifa, School  
of Public Health 199 Aba Khoushy Ave. Mount Carmel, Haifa 3498838, Israel

\* Corresponding author. E-mail: davidmanheim@gmail.com

**Conflict of interest and funding:** Authors report no funding and no conflicts of  
interest.

**Human Subjects Approval:** An IRB determined the data collection and  
analysis does not constitute human subjects research, and was approved.

**Data availability statement:** Code is available on GitHub, per the appendix.  
Data is available from twitter based on available files on GitHub, or per twitter  
terms and conditions, full dataset is available for academic use on request.

**Abstract:**

This paper considers how health education organizations in the World Health Organization's Vaccine Safety Network (VSN) use Twitter to communicate about vaccines with the public, and whether they answer questions and engage in conversations.

Almost no research in public health, to our knowledge, has explored conversational structure on social media among posts sent by different accounts. Starting with 1,017,176 tweets by relevant users, we constructed two corpuses of multi-tweet conversations. The first was 1,814 conversations that included VSN members directly, while the second was 2,283 conversations mentioning vaccines or vaccine denialism. The tweets and user metadata was then analyzed using an adaptation of Rhetorical Structure Theory.

In the studied data, VSN members tweeted 12,677 times within conversations, compared to their 37,587 lone tweets. Their conversations were shorter than those in the comparison corpus ( $P < 0.0001$ ), and they were involved in fewer multilogues ( $P < 0.0001$ ). We also see that while there is diversity among organizations, most were tied to the pre-social-media broadcast model. In the future, they should try to converse more, rather than tweet more, and embrace best-practices in risk-communication.

**Keywords:** risk communication; twitter; social media; conversation; computational linguistics; engagement

## **Introduction**

Social Media platforms are increasingly important platforms for public communication, and Twitter in particular is being used extensively in various positive capacities in public health (Denecke et al, 2013; George et al, 2013; Paul and Dredze, 2011). Social media has become important both as a platform for organizations attempting to spread important public health messages (Gesser-Edelsburg and Shir-Raz, 2016), for engaging with the public (Chou et al., 2013), and for understanding response to those messages. The 2019 Digital Media Report found that 25% of Americans mainly discover news via social media, and over 20% of adults in a selection of western countries use Twitter each week (Newman, 2019). Not only is this significant, but the proportion of the population using and seeking information from newer platforms seems likely to continue to grow.

The World Health Organization (WHO) is an international organization that provides leadership and builds partnerships in international health, and along with other key functions, helps set standards and norms for the international community (WHO, 2007). One such partnership building and norm setting program run by the World Health Organization is the Vaccine Safety Network (VSN), “a global network of websites... that provides reliable information on vaccine safety,” with goals that include “communicating vaccine safety information through a diversity of digital channels” (WHO, 2019). Both the Vaccine Safety Network and many of the organizations in the network, have twitter accounts, which were studied. Risk communication is a core need for such organizations, (Gesser-Edelsberg et al., 2015; Dickmann et al, 2016) and this network provides an excellent starting point for looking at such communication on Twitter related to vaccines.

Much previous research on twitter for public health focuses on the lay public (Bricker, 2019), or experts, rather than organizations. There have been calls to focus on engagement with the public (Linan, 2019), but much of the analysis is on either “following” / “follower” relationships, as in Cavazos-Rehg et al. (2014), or tweet contents (sentiment analysis, etc.), rather than conversation - tweets, replies, and responses. Some studies have considered conversation, at least at smaller scales. Lakon et al. (2016) staged an intervention which looked at conversation structure for small groups of smokers with up to four participants.

The current work focuses on engagement, and looks at the other half of risk communication, how organizations engage with experts and the lay public. This is particularly important in the polyvocal context faced by health organizations (Gesser-Edelsburg, 2016). The discussion of MMR and measles which the data collected for this analysis focuses on poses particular challenges, noted by Bricker et al. (2019), which require effective health communication strategies (Vaughan and Tinker, 2009).

Communication depends on the medium being used, and conversation on twitter differs from other types of expert and health education interactions, even compared to other online platforms (Paul and Dredze, 2011; Gesser-Edelsberg and Shir-Raz, 2016). It is a vehicle for experts to communicate both with each other, and with the public (Gesser-Edelsberg et al., 2016; Lovejoy and Saxton, 2012), not just broadcast messages. This is critical because conversations, rather than the validity or convincingness of the arguments made, is particularly critical for influencing the public (George et al, 2013).

Conversation and outreach are also critical given declining trust generally. Less than half of people say they trust even the news institutions that they prefer using (Newman, 2019). The issue is particularly critical for measles and conversations about vaccines. In the last several years, large outbreaks in the Netherlands, Sweden, Italy,

Japan, and the United States have each occurred, all in part due to low vaccination rates (Bricker, 2019; CDC, 2019). The recent outbreaks in the western world stem in large part from vaccine hesitancy (WHO, 2018). For the MMR vaccine specifically, anti-vaccine sentiment has created gaps in vaccine coverage that led to a spike in cases. (Bricker, 2019)

Institutions attempting to provide leadership in public health and communication are already attempting to address the lack of trust, and engagement is a key strategy for doing so. Given the above issues of trust in general, and vaccine skepticism specifically, the importance of reaching the public's concerns and engaging them, not just broadcasting messages, should be clear. However, changing minds requires interaction, not just exposure, or social media engagement as typically measured (Bruns & Stieglitz, 2013; Keller et al., 2014; Chretien et al. 2015). Given the different goals for twitter usage in medicine and public health, there are different ways of engaging with different advantages and risks, which can be best accomplished using different types of discourse (Keller et al., 2014). These will also involve different conversational structures, which this research attempts to analyze.

This paper is the first example in infectious disease research, to our knowledge, that explores the structure of discourse on social media on a large scale by considering the relationship between multiple posts sent by different accounts. The analysis focuses on discussions between public health organizations, experts, and the general public about infectious disease and vaccines, in order to understand conversation structure in historical Twitter data. The analysis considers key risk-communication goals and whether they are being met. In doing so, it applies and validates a methodology for analyzing discourse among users on Social Media with a novel and widely applicable, if data- and time-intensive, sampling strategy.

Previous work in this vein includes Neiger et al. (2013), which considered local health departments' use of Twitter. In addition to sentiment analysis, they considered whether they are attempting outreach or promotion, albeit by analyzing tweet contents rather than conversational structure itself. More generally, Lovejoy and Saxton (2012) looked at nonprofits. While they considered conversation (or "dialogue") it was only based on whether organizations used the "reply" feature.

### **Materials and Methods**

Below we provide an outline of the methods used for data analysis. Further details are available in the appendix, and the GitHub repository.

#### ***Sample***

An initial sample was retrieved that included all retrievable tweets sent during the calendar year 2018 by accounts belonging to the 52 organizations in the WHO's Vaccine Safety Network with twitter accounts. This included 35,878 tweets, of which 11,454 were retweets of other accounts, and 3,869 were non-retweeted interactions with other users.

This sample was used to find tweet threads for the subset which involved interaction - either replies or quote-tweets - by the organization. Any tweets by accounts identified as "Bots" were excluded. This sampling strategy allows extraction of conversations, rather than typical methods of retrieving tweets that use a hashtag or key term unrelated to the conversational structure. This is conceptually similar to Cogan et al's (2012) early "complete conversation" approach, but is not hobbled by Twitter's more recent 1% sample limit.

After retrieval, we constructed two datasets that allowed us to focus on the question of how organizations were using Twitter, and compare their usage to a related

baseline of conversations. The first dataset was conversations involving the VSN Twitter accounts. The second was all conversations where at least one tweet in the conversation contained a keyword related to measles or vaccinations, described in the appendix. The final analyzed corpus included 3,998 threads with a total of 24,192 tweets. This again excludes tweets from accounts identified as bots, discussed further in the appendix.

### *Heuristic Classification Methods*

Because of the size of the dataset and number of accounts in the dataset, it is infeasible to manually classify the accounts. In place of manual classification, a set of heuristics was developed and applied, then manually reviewed to assess accuracy. The two heuristics described in the appendix were (1) automated analysis of the web sites associated with the account, and (2) keyword filters applied to account descriptions. Note that for non-verified accounts, which are the majority of accounts, neither the description nor the web sites are verified by Twitter. This means that classification is based on how accounts represent themselves, and is what other users see when considering the source, and therefore the authenticity or trustworthiness of a tweet. Unfortunately, twitter can only provide user data at present, not the data that was extant at the time of the tweet being sent.

### *Identifying Deceptive Activities and “Bots”*

A significant concern in analyzing twitter conversations is whether the accounts were run by bona-fide users, or “bots” - a somewhat conceptually unclear category. This is especially true for vaccines, where there has been concerns voiced about astro-turfing, and evidence that governments attempted to push disinformation. Supporting this concern, there is research by Boniatowski et al. (2018) showing that bots, including

those involved in government disinformation campaigns, have been used. While it is unfortunately impossible to reliably identify all accounts that are being used for nefarious purposes, we used the “Bot or Not” classifier (Yang et al, 2019) to consider whether the accounts were, in fact, bots, and report the scores, as well as removing detected bots from the analysis. Bot detection is an ongoing struggle, and as it improves, bots adopt new techniques for remaining unidentified.

### ***Theoretical Framework***

In this paper, we build on Rhetorical Structure Theory (RST), a formal model introduced by Mann and Thompson (1988) for systematic analysis of relationships between different text spans, which has been suggested can be applied to twitter (Sidarenka et al, 2015). RST provides a topology of conversation and interaction, which is extended here to allow identification of conversational flow between multiple speakers. This is distinct from the more widely explored topology of follower networks (Kwak et al, 2010), or conversation analyses where conversation is defined by shared hashtag usage (Derczynski, 2013). This novel use of RST allows us to consider how health organizations engage with doctors, experts, and the public, both in general, and in the context of recent measles outbreaks and discussions of the measles vaccine. Critically, we can explore whether organizational twitter use is primarily one-way communication, broadcasting information without responding to or interacting with other users, or if there is conversational interaction. In addition, when conversation occurs we explore differences between intra-expert and expert-organization discussions versus engaging with the broader public either responsively, or interactively. Where it is interactive, we further characterize the types of interaction, as explained below and in the methodological Appendix.



Note that retweets are not reflected in the conversational structure we consider using RST, since the retweeting account does not itself make a statement. This could be modified by assigning multiple authors to a retweeted statement, but this is misleading given that, as the standard disclaimer notes, retweets are not endorsements.

### ***Research Approach***

The purpose of this analysis is to indicate whether the organizations were involved in one-way (broadcast-type) communication, whether they replied but did not engage, or whether there was the type of more intensive communication that social media allows. In the cases that there was communication, we wish to understand the form of that conversation.

The analysis performed in this paper focuses on conversation, and does not involve content analysis. While traditional content analyses are valuable, such methods struggle to capture context and interaction (Smith & Gallicano, 2015). In contrast, our approach is particularly appropriate for understanding whether and how organizations engaged with the public, especially given large and massively multi-user datasets.

In the first stage we defined engagement. The term is used differently in public health, in risk communication, and in social media. In public health, engagement can mean participation in an intervention or interacting with health providers. In social media it is interaction with the content with the user interface - expanding the tweet, liking it, retweeting, or replying. But Smith and Gallicano (2015) point out the critical difference between interaction with content, mental engagement with the topic, and having an interactive discussion which is critical for risk communication. We are focused on this last form of engagement, which we call conversation. Unfortunately, typical metrics, such as likes and retweets, measure success at broadcasting, and are

indications of reach and agreement rather than indicating engagement relevant for risk communication, i.e. conversation.

In the second stage, we applied the definition to build a conceptual framework based on RST (Mann & Thompson, 1988), a descriptive theory from linguistics, for understanding conversation on twitter to measure engagement. A preliminary analysis considered organizations' individual tweets to evaluate whether organizations use Twitter for conversation at all, either by responding to or quoting tweets by other users, or if Twitter was used only for tweeted announcements or dissemination of information from trusted sources by linking or retweeting. Length of conversation is a key factor, but so is structure. Where organizations engaged in any form of interaction, we analyzed these conversations with the adapted RST framework. The analysis categorized conversations into monologues, where the account replied only to itself; reply conversations, where replies were received but no dialogue occurred; dialogues, where two or more users replied to one another; and multilogues, where more than two users responded within a single discussion.

This analysis is critical because interaction alone is not necessarily sufficient. The typology above therefore differentiates between responsive dialogue, replying to specific tweets, or more intense conversation between multiple users, experts, and/or organizations. Conversation is a critical part of health communication according to Park et al. (2016), and the analysis builds on the approach of Lovejoy and Saxton (2012), among others.

The value of conversation for risk communication also depends on participants. Gesser-Edelsburg and Shir-Raz (2016) identify several challenges in creating credible discourse on online platforms, which include: “public health experts versus ‘people with agendas’; conflicts of interest; [and] facts/rationality v. emotions/myths”. To overcome

these challenges, engagement with the public should be done by clearly identified experts and well-known organizations pointing to verifiable and trustworthy sources. Clear identification can be measured by looking at how experts identify themselves, as discussed below, and reference to trustworthy sources can be considered by looking at inclusion of links in replies, and the sources referenced.

Given this, the third and final stage considered who was participating in conversations. In addition to the 52 organizations of the WHO's list, there were a variety of other organizations, and individuals that replied, were replied to, or were quoted, and therefore appear in the dataset. The organizations include both healthcare organizations such as hospitals, and non-profits, as well as other organization types including news, government bodies, and others. The individuals range from experts, to the lay public, to celebrities and other noteworthy personalities such as leaders of key organizations and news reporters. Unfortunately, as discussed below and in the appendix, it is often difficult to know which is which, but heuristic methods were used to perform classification. Given this identification, it is possible to see whether conversations were in fact engaging with the public.

To consider this, we present a breakdown of a number of characteristics, including thread type, length, number of participants, and participant types - expert, organizational, or member of the public. User types were based on a heuristic method, details of which are available in the Appendix.

## **Results**

The data retrieved was a total of 1,017,176 tweets across the participants in conversations with WST accounts. This was filtered and partitioned into two partially overlapping datasets. The first, the organization dataset, was all retrieved conversations

directly involving the organizations. This was a total of 1,826 conversations involving 6,630 tweets.

The second, the relevant conversations dataset, were all conversations captured in the tweet set in which a vaccine relevant keyword (defined in the methodology appendix,) was used somewhere in the extended conversation, by any user. This was a total of 2,427 conversations, with a total of 24,192 tweets. 255 of the conversations, with 1,257 tweets, were included in both sets.

Conversational length, in number of tweets, is a simple measure of engagement, as noted above. A Welch 2 sample, 1-sided t-test shows that the VSN conversations were statistically significantly shorter ( $t = 7.8984$ ,  $df = 2433.7$ ,  $p < 0.0001$ ), with a mean of 3.6 tweets in VSN conversations, compared to a mean of 9.5 tweets for relevant conversations. The distribution of the conversation sizes as a percentage of the total is shown below.

[Image 1 - Conversation Sizes]

### ***Tweet Types***

On Twitter, tweets can be replies, retweets, or tweets which are none of these, which we call isolated tweets. They can also be both replies and retweets. For example, a reply tweet can quote a third tweet, or the tweet it is replying to. In the VSN dataset, we find that VSN members tweeted or retweeted others a total of 37,587 times. Of these, 7,479 were replies, 5,053 were quote-tweets, and a further 145 were replies that also quoted a tweet. There was significant variance between accounts, however. For instance, as of the date of retrieval, the total number of tweets sent by the accounts, including those from before the retrieval window, ranged from 24,809 tweets sent, to only 42.

[Image 2 – Overall VSN Tweet Distribution]

[Image 3 – Small Multiples]

Note that while most members rarely replied, some did so heavily. Similarly, only a few accounts ever both replied and quote-tweeted in a single tweet (shown in red,) but one did so relatively often.

### *Analysis of Conversation Types*

The VSN conversations had a significantly smaller ( $X^2 = 154.93$ ,  $df = 1$ ,  $p\text{-value} < 0.0001$ ) proportion of multilogues, 6% compared to 19%. Given the number of relevant conversations, and the number of reply conversations, this was not for lack of opportunity.

[Table - Conversation Categories]

The breakdown of these conversations by the user types involved is shown below.

[Image 3 – Sankey Diagram]

The two datasets display a clear difference in the number of multilogues. This is similar to the previous observation, and we see that VSN members engaged in 14% fewer multilogues than occurred in relevant conversations.

### **Question and Answer Frequency**

In the datasets, we also looked at questions asked. The table below compares the two corpuses analyzed, and shows that VSN members responded to questions in conversations infrequently.

[Table – Conversations and Questions]

On other social media services, there is work showing that questions are common, and answers by experts are sought out. Unfortunately, the organizations best

placed to be reliable sources of that information seem not to engage in those conversations; less than half of all conversations with questions had answers by VSN members. In the relevant conversations dataset, most of the answers, 3,450, were from users that were not identifiable organizations or experts. Given the focus on providing information, it seems unfortunate that these organizations are often unresponsive when answers are sought. Perhaps even worse, if it is clear that organizations don't typically reply to questions, the public will instead seek answers from sources which may not provide authoritative or even correct responses.

## **Discussion**

The organizations whose twitter use is analyzed in this paper are critical parts of the health education ecosystem, according to the World Health Organization (2019), but their use of Twitter in this context leaves much to be desired.

We looked for conversation and interaction between health organizations, experts, and the public. These did occur, but vastly outnumbered by broadcast-focused use of twitter. As noted, VSN members engaged in 14% (12% to 16%,  $P < 0.0001$ ) fewer multilogues than occurred in relevant conversations, and answered fewer than half of the questions asked in conversations they participated in.

The organizations mostly use twitter to make announcements. Provision of reliable health information is critical, and these organizations continue to carry out that mission well. Still, many fail to capitalize on social media as a way to interact with the public. As the World Health Organization (2020) notes, a key advantage of social media is allowing people to “engage in conversations,” which requires goals beyond “using more visuals, pictures and infographics to simplify information, tell better stories, and create a fast but lasting impact.”

The analysis in this paper validates the use of conversation analysis in understanding usage of twitter as a social media platform more generally by showing that material differences in usage exist between accounts and between corpuses. Previous work has highlighted the importance of conversation, but large scale studies used far less direct methods for their analyses based on hashtags and counting replies. Given the existence of this new approach for more in depth large scale analysis, the method should be considered as an alternative or complement to content analysis when considering how social media is used and experienced.

The analysis also gives insight about how to better address the recommendations for creating engagement; focus more on replying to those who respond and ask questions, in addition to focusing on using hashtags and images, or focusing on follower counts and retweets. It would be concerning if metrics for replies were to completely crowd out current metrics for social media use in organizations, but a more comprehensive approach that included different metrics could certainly encourage a balance.

(Manheim, 2018; Sugimoto et al 2017)

A secondary point that emerges from this analysis is the way in which health organizations' lack of coordinated response and engagement with the public has left individual experts with the burden of addressing these issues. While domain experts have performed admirably, and even large parts of the broader public have been helpful in spreading useful health education messages, they are filling in a role that health organizations are not.

The method used has two key limitations, discussed more completely in the Appendix. The first is that retrieval is limited to identified users, which is critical when understanding failure to reply. In the current case, however, the captured replies are

sufficient to show the noted lack of engagement.

The second is that analysis of participants required for conversation analysis relies on effective identification of accounts, and this is challenging to do at scale. However, individual experts whose accounts are not clearly identified, for the purposes of public understanding, not experts. In addition, for organization identification the set of health education organizations analyzed was specified before data collection began, so that the impact was more limited. Still, better methods are required for more complete analyses.

## **Conclusion**

There is a clear difference between the conversations on Twitter that health organizations are involved in and the conversations which discuss vaccines and measles. It seems clear that most organizations are still using Twitter as a one-way communication channel. This can be addressed by reconsidering the communication strategies used on Twitter, and dedicating people and resources to interaction. This goes farther than tweeting more, or even replying and conversing more. Risk communication requires considering of the context of questions. This will require engagement with the public that extends beyond health education organizations' role as dispassionate providers of truth, to greater real engagement with the public

Despite the great importance that health organizations attach to the use of social media, and despite existing guidelines and accumulated knowledge at the application level, there are still gaps. If organizations want to begin allocating appropriate resources, they can begin to create a continuous conversation with the public. As current trends in misinformation and infodemics make clear, the need is critical, and not yet fully addressed. Improving social media risk communication is not a one-time effort restricted to a single class of misinformation or acute event, but is part of a continuing



need in public health.

## Bibliography

- Denecke K, Krieck M, Otrusina L, Smrz P, Dolog P, Nejd W, Velasco E. How to exploit twitter for public health monitoring? *Methods Inf Med.* 2013;52(04):326–339.
- George DR, Rovniak LS, Kraschnewski JL. Dangers and opportunities for social media in medicine. *Clin Obstet Gynecol.* 2013;56(3).
- Paul MJ, Dredze M. You are what you tweet: Analyzing twitter for public health. In: *Fifth International AAAI Conference on Weblogs and Social Media.* 2011.
- Gesser-Edelsburg A, Shir-Raz Y. *Risk Communication and Infectious Diseases in an Age of Digital Media* [Internet]. 1st ed. Routledge; 2016 [cited 2019 Jul 28]. 188 p. (Routledge Studies in Public Health). Available from: <https://www.routledge.com/Risk-Communication-and-Infectious-Diseases-in-an-Age-of-Digital-Media/Gesser-Edelsburg-Shir-Raz/p/book/9781138186064>
- Chou WY, Prestin A, Lyons C, Wen KY. Web 2.0 for health promotion: reviewing the current evidence. *American journal of public health.* 2013 Jan;103(1):e9-18.
- Newman N. *Digital News Report 2019* [Internet]. Reuters Institute for the Study of Journalism, University of Oxford; 2019 [cited 2019 Jul 28] p. 153. Available from: [https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2019-06/DNR\\_2019\\_FINAL\\_1.pdf](https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2019-06/DNR_2019_FINAL_1.pdf)
- World Health Organization. *Working for health: an introduction to the World Health Organization.* 2007 [cited 2019 Jul 28]; Available from: [https://www.who.int/about/brochure\\_en.pdf?ua=1](https://www.who.int/about/brochure_en.pdf?ua=1)
- World Health Organization. *About Vaccine Safety Net.* 2019 Mar 18; Available from: <https://www.vaccinesafetynet.org/vsn/vaccine-safety-net>
- Gesser-Edelsburg A, Shir-Raz Y, Walter N, Mordini E, Dimitriou D, James JJ, Green MS. The public sphere in emerging infectious disease communication: recipient or active and vocal partner? *Disaster Med Public Health Prep.* 2015;9(4):447–58.
- Dickmann P, Abraham T, Sarkar S, Wysocki P et al.. Risk Communication as a Core Public Health Competence in Infectious Disease Management: Development of the ECDC Training Curriculum and Programme. *Eurosurveillance* 21:14
- Bricker B. A Case Study of MMR Vaccine Skeptics. *Recover Argum.* 2019;16.
- Linan BS. Communicating with the public is key to public health. *Amer Journal Pub Health.* 2019 Mar;109(4), pp. 548–549
- Cavazos-Rehg P, Krauss M, Grucza R, Bierut L. Characterizing the Followers and Tweets of a Marijuana-Focused Twitter Handle. *J Med Internet Res.* 2014 Jun 27;16(6):e157.
- Lakon CM, Pechmann C, Wang C, Pan L, Delucchi K, Prochaska JJ. Mapping engagement in Twitter-based support networks for adult smoking cessation. *American journal of public health.* 2016 Aug;106(8):1374-80.
- Vaughan E, Tinker T, 2009: Effective Health Risk Communication About Pandemic Influenza for Vulnerable Populations. *American Journal of Public Health* 99,

- Lovejoy K, Saxton GD. Information, community, and action: How nonprofit organizations use social media. *J Comput-Mediat Commun.* 2012;17(3):337–53.
- Center for Disease Control,. Measles Cases and Outbreaks [Internet]. 2019 [cited 2019 Aug 1]. Available from: <https://www.cdc.gov/measles/cases-outbreaks.html>
- World Health Organization. Measles cases spike globally due to gaps in vaccination coverage [Internet]. World Health Organization News Room. 2018 [cited 2019 Jul 28]. Available from: <https://www.who.int/news-room/detail/29-11-2018-measles-cases-spike-globally-due-to-gaps-in-vaccination-coverage>
- Bruns A, Stieglitz S. Towards more systematic Twitter analysis: metrics for tweeting activities. *Int J Soc Res Methodol.* 2013;16(2):91–108.
- Keller B, Labrique A, Jain KM, Pekosz A, Levine O. Mind the gap: social media engagement by public health researchers. *J Med Internet Res.* 2014;16(1):e8.
- Chretien KC, Tuck MG, Simon M, Singh LO, Kind T. A digital ethnography of medical students who use Twitter for professional development. *J Gen Intern Med.* 2015;30(11):1673–80.
- Neiger BL, Thackeray R, Burton SH, Thackeray CR, Reese JH. Use of Twitter Among Local Health Departments: An Analysis of Information Sharing, Engagement, and Action. *J Med Internet Res.* 2013;15(8):e177.
- Cogan et al., Reconstruction and Analysis of Twitter Conversation Graphs. Proceedings of the First ACM International Workshop on Hot Topics on Interdisciplinary Social Networks Research
- Broniatowski DA, Jamison AM, Qi S, AlKulaib L, Chen T, Benton A, Quinn SC, Dredze M. Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate. *Am J Public Health.* 2018 Aug;108(10):1378–1384.
- Yang K, Varol O, Davis CA, Ferrara E, Flammini A, Menczer F. Arming the public with artificial intelligence to counter social bots. *Hum Behav Emerg Technol.* 2019;1(1):48–61.
- Mann WC, Thompson SA. Rhetorical structure theory: Toward a functional theory of text organization. *Text-Interdiscip J Study Discourse.* 1988;8(3):243–81.
- Sidarenka U, Bisping M, Stede M. Applying Rhetorical Structure Theory to twitter conversations. In Proceedings of the Workshop on Identification and Annotation of Discourse Relations in Spoken Language (DiSpol) 2015. p. 1–2.
- Kwak H, Lee C, Park H, Moon S. What is Twitter, a social network or a news media? In *AcM*; 2010. p. 591–600.
- Derczynski LR, Yang B, Jensen CS. Towards context-aware search and analysis on social media data. In *ACM*; 2013. p. 137–42.
- Smith BG, Gallicano TD. Terms of engagement: Analyzing public engagement with organizations through social media. *Comput Hum Behav.* 2015;53:82–90.
- Park H, Reber BH, Chon M-G. Tweeting as health communication: health organizations' use of Twitter for health promotion and public engagement. *J*

Health Commun. 2016;21(2):188–98.

World Health Organization, Communicating for Health: Social Media.

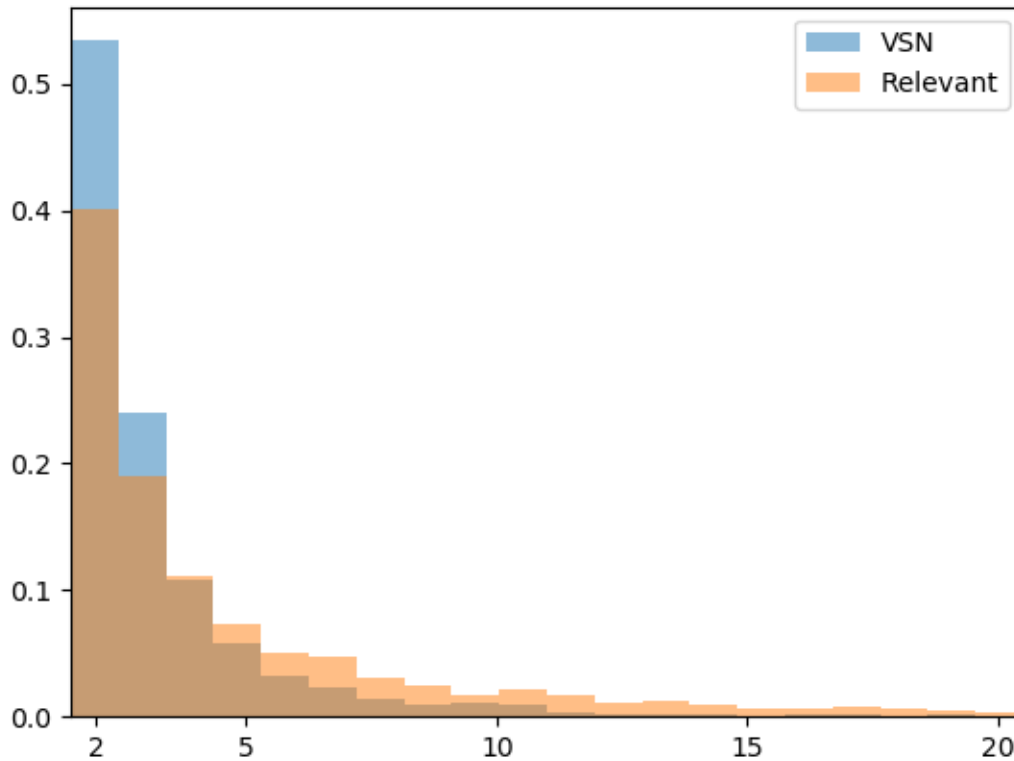
<https://www.who.int/communicating-for-health/functions/social-media/en/>

Manheim D, “Building Less Flawed Metrics: Dodging Goodhart and Campbell’s Laws.” 2018. Available at: <https://mpira.ub.uni-muenchen.de/98288/>

Sugimoto CR, Work S, Larivière V, Haustein S. Scholarly use of social media and altmetrics: A review of the literature. *J Assoc Inf Sci Technol.* 2017;68(9):2037–62.

# Images

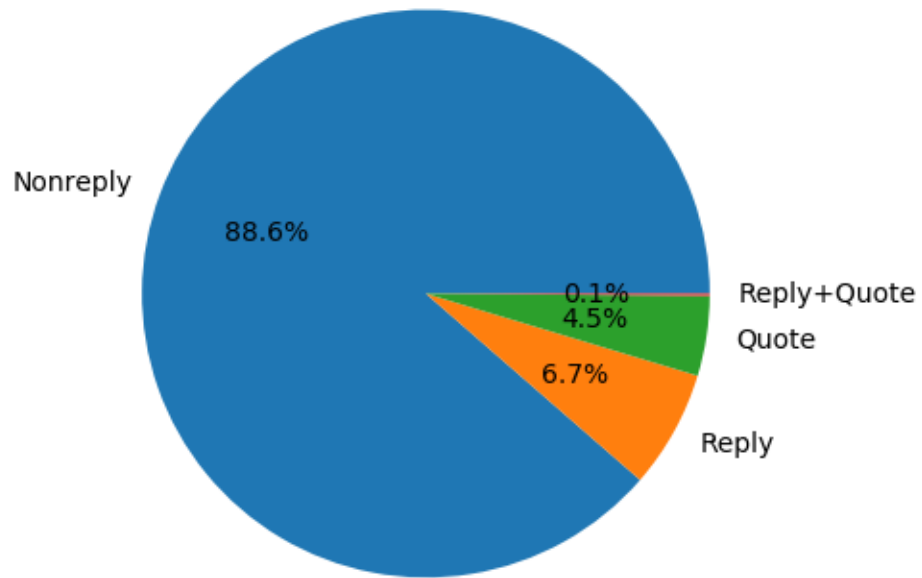
Image 1 – Conversation Sizes



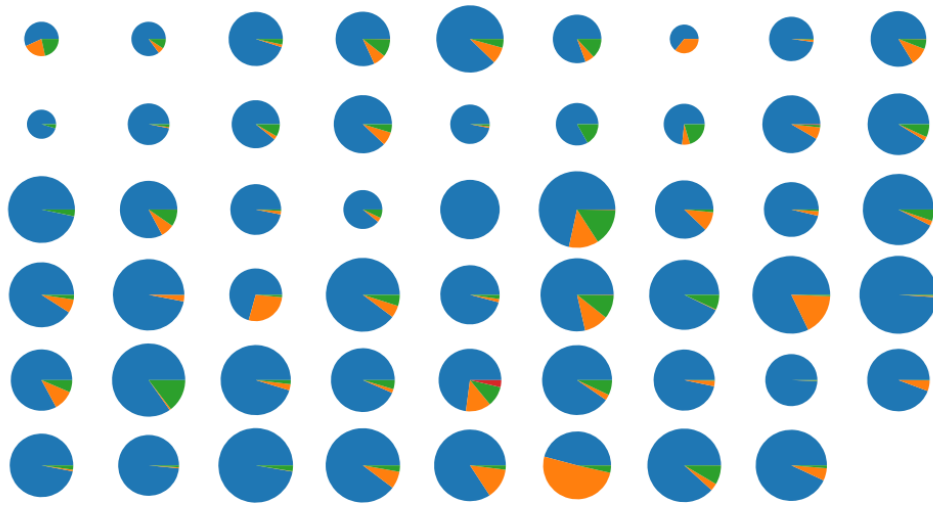
## Caption:

The majority of conversations by Vaccine Safety Network members were a single reply, and the conversation length, mean 3.6 tweets, was significantly shorter ( $t = 7.8984$ ,  $df = 2433.7$ ,  $p < 0.0001$ ) than those in the corpus of relevant conversations, mean 9.5 tweets. Note that the graph is truncated at 20 tweets in a conversation, and single tweets are excluded, since they are not conversations. The VSN dataset has under 1% of tweets longer than 20, with the longest being 116 tweets, while the Relevant dataset has just under 10% of conversations longer than 20 tweets, with the longest being over 1,400 tweets.

Image 2 – Overall VSN Tweet Distribution

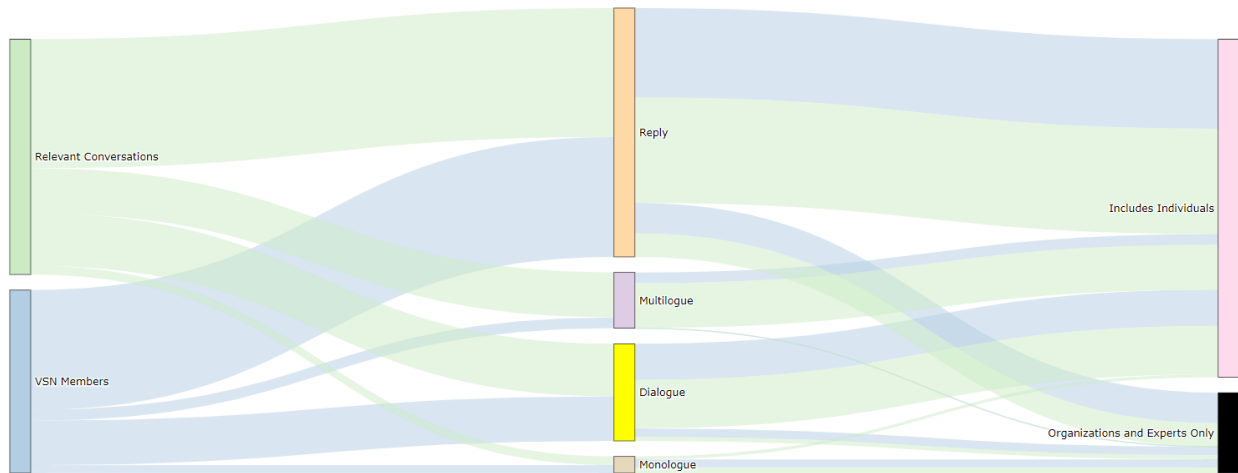


Caption: This graphs the distributions of tweet types in the VSN conversations dataset. This excludes isolated tweets, which as noted, formed the vast majority of all tweets, so that "nonreply" are tweets which are part of a conversation that follows the tweet.



Caption: This is a small multiples visualization of per member graphs, with sizes scaled logarithmically by the total number of tweets sent by each account, (not just those within the retrieval window, which is capped at 3,200), so the most active account is approximately triple the size of the least active one.

Image 3 – Sankey Diagram



Caption: This shows the proportion of each type of conversation (Middle Column) in each dataset (Left Column). The VSN members' multilogues were less frequent, containing 14 percentage points fewer such (95% CI, 12-16%, N-1 Chi-squared test,  $P < 0.0001$ ) than the relevant conversations dataset. The color-coded subsets of each type which include individuals and unclassified users, rather than just organizations, are then shown. Given the nature of conversations and participants, multilogues almost all include at least some individuals, while monologues contain almost none.



Table - Conversation Categories

	VSN Conversation Corpus	Relevant Conversation Corpus
Total Conversations	1814	2427
Users Involved	2199	4283
Monologues	79	86
Reply-Conversations	1197	1308
Dialogues	442	546
Multilogues	108	487

The VSN dataset contained 6% multilogues, a significantly smaller proportion ( $X^2 = 154.93$ ,  $df = 1$ ,  $p\text{-value} < 0.0001$ ) than the relevant conversations dataset, which was 19% multilogues.

Table – Conversations and Questions

	VSN	Relevant
Total Conversations	1814	2427
Total Questions	2050	4972
Questions by VSN Members	367	54
Conversations with Questions	1176	1372
Answers	1229	5037
Answers by VSN Members	557	115

Table Caption: Questions were found using the method of Li et al, which found that the simple heuristic of detecting question marks in tweets found almost 85% of questions, with precision of 97%. Applying that method to our datasets, it is possible to identify not only questions, but the answers, i.e. other users' tweets that replied to the question.