“Yeah, this graph doesn’t show that”: Analysis of Online Engagement with Misleading Data Visualizations

Maxim Lisnic
maxim.lisnic@utah.edu
University of Utah
Salt Lake City, Utah, USA

Alexander Lex
alex@sci.utah.edu
University of Utah
Salt Lake City, Utah, USA

Marina Kogan
kogan@cs.utah.edu
University of Utah
Salt Lake City, Utah, USA

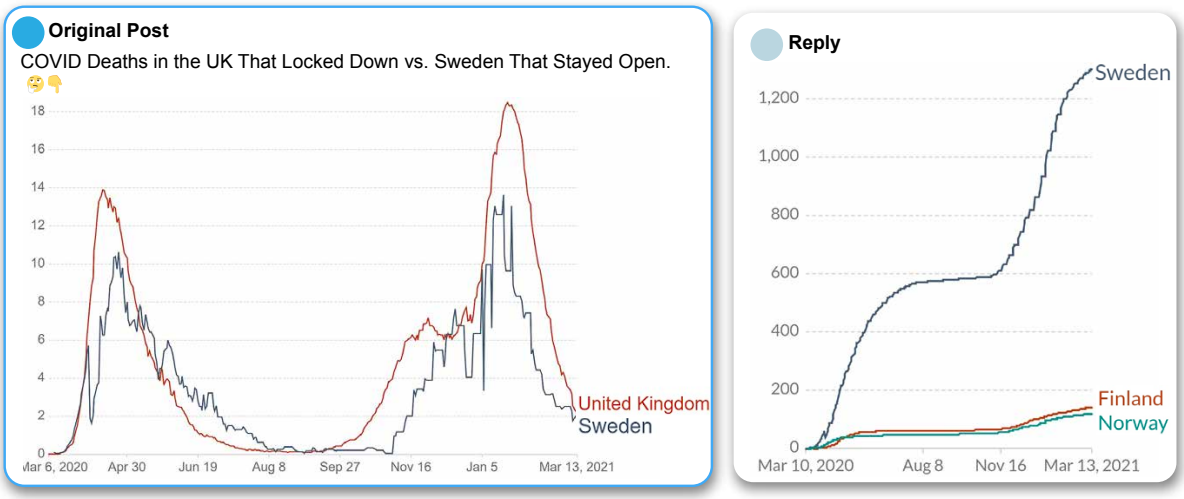

Figure 1: An example of a post with a data visualization-supported insight and an associated reply. The author of the original post shares a chart comparing COVID-19 deaths in the UK and Sweden, making an argument that lockdowns are not effective because the case curves of the two countries are similar. The reply makes a counter-argument (without the use of text) by sharing an analogous data visualization that supports the opposite conclusion, which shows that Sweden had many more cases than Finland and Norway, comparable Nordic countries.

ABSTRACT

Attempting to make sense of a phenomenon or crisis, social media users often share data visualizations and interpretations that can be erroneous or misleading. Prior work has studied how data visualizations can mislead, but do misleading visualizations reach a broad social media audience? And if so, do users amplify or challenge misleading interpretations? To answer these questions, we conducted a mixed-methods analysis of the public’s engagement with data visualization posts about COVID-19 on Twitter. Compared to posts with accurate visual insights, our results show that posts with misleading visualizations garner more replies in which the audiences point out nuanced fallacies and caveats in data interpretations. Based on the results of our thematic analysis of engagement, we identify and discuss important opportunities and limitations to effectively leveraging crowdsourced assessments to address *data-driven misinformation*.

CHI ’24, May 11–16, 2024, Honolulu, HI, USA

CCS CONCEPTS

• **Human-centered computing** → **Visualization theory, concepts and paradigms**; **Empirical studies in collaborative and social computing**.

KEYWORDS

Visualization, social media, misinformation, COVID-19

ACM Reference Format:
Maxim Lisnic, Alexander Lex, and Marina Kogan. 2024. “Yeah, this graph doesn’t show that”: Analysis of Online Engagement with Misleading Data Visualizations. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI ’24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3613904.3642448

1 INTRODUCTION

Advances in data collection and data literacy and the rapid spread of information on social media have enabled us to use data visualizations to quickly discover and spread awareness about signs of otherwise invisible phenomena, such as climate change or viral disease epidemics. However, although data helps us uncover evidence of an event or make sense of it, an erroneous analysis may provide an illusion of evidence, lead to false discoveries or

false accusations, or trigger rumors. Whether intentional or stemming from misunderstanding, incorrect or incomplete interpretations of visualizations on controversial topics have the potential to cause harm by spreading misinformation. Indeed, research has documented that misleading data visualizations have been used in support of misinformation on a variety of topics, such as COVID-19 skepticism [23, 28], climate change denial [51], false claims of election fraud [30], and QAnon conspiracy theories [12].

Prior work has highlighted the ways in which data visualizations can deceive an audience due to visual tricks and mirages [7, 29, 32, 36]. However, charts that support misinformation arguments are most commonly well-designed and mislead viewers by being vulnerable to biased framing, misinterpretations, and logical fallacies. [28]. With the rise in popularity of interactive data exploration sites for COVID-19 data such as OurWorldInData [31] or Worldometer [49], the ability to create professional-looking data visualizations has become more democratized and accessible to non-expert users. Consequently, however, the problem of well-designed charts being vulnerable to misinterpretations has reached the scale of mass audiences and is used to fuel misinformation arguments on social media, with 42% of COVID-19-skeptic visualizations shared on Twitter[1] being screenshots of reputable data explorers with a recognizable style and branding [28]. For instance, the original post in Figure 1 attempts to promote a COVID-skeptic argument by sharing a data explorer chart showing a cherry-picked data selection, which was effectively countered by the analogous visualization in the reply.

However, can social media audiences always reliably point out such misleading tactics? Despite the fact that numerous studies have examined the spread [42], correction [2, 3, 46], and moderation [37, 52] of misinformation on social media, this research is mostly focused on text and has yet to examine how people share and react to visualization-supported misinformation. As a consequence, it is unclear whether existing findings on misinformation interventions also apply to misinformation supported by misleading visualizations. In their recent work, Weikmann and Lecheler discuss that visual disinformation, including misleading visualizations, is "its own type of falsehood [that] differs from textual disinformation" because it allows for a higher level of manipulative sophistication [48]. All of the above points to the existence of a research gap in understanding the public engagement with, and the potential for mitigation of, data visualization-supported misinformation that opens the door for harmful rumors and conspiracies.

Our paper attempts to fill this research gap by presenting the results of a mixed-methods study of engagement with both misleading and accurate insights in COVID-19 data visualization posts on Twitter. We attempt to answer the following questions:

> **RQ1**: Do misleading insights in a data visualization post have an effect on the count and duration of its engagement?
>
> **RQ2**: Do people identify and raise awareness about misleading data insights in their responses?

Based on the results of our work and a review of existing misinformation literature, we discuss the ways in which *data-driven misinformation* in visualization posts is distinct from factual forms of misinformation that are typically studied, such as misinformation based on text or deepfakes. We posit that **existing mitigation strategies may not be sufficient in supporting the verification of nuanced misinformative data interpretations** such as statistical fallacies or data collection caveats. Moreover, data visualizations are associated with credibility indicators that are distinct from those that apply to other types of misinformation, namely the source of the chart and the data, perceived data literacy and analytical expertise of the author, and perceived data integrity.

This paper makes several contributions:

- Firstly, we conduct a quantitative study of engagement with posts containing data visualizations on social media. Our results show that posts offering interpretations of data are shared twice as frequently—regardless of the their accuracy. Misleading data interpretations garner an additional 60% more replies compared to accurate insights.
- Secondly, we present the results of a thematic analysis of replies to posts with interpretations of data visualizations through a series of case studies. Our findings show that the crowd has the potential to find and reason about nuanced caveats in misleading data-driven insights on social media.
- Thirdly, the results of our thematic analysis also describe important limitations of the crowd's ability to effectively verify misleading data-driven insights using the existing platform affordances. We discuss approaches that could help tackle these limitations, such as meta-analyses, counter-analyses, and trust-building for data sources and analysts.
- Lastly, we describe the differences between data-driven misinformation and other forms of misinformation on social media and discuss important considerations in designing interventions to address it.

2 RELATED WORK

In this section, we discuss how existing work on misleadingness of visualizations and recent studies of online misinformation point to the existence of a research gap in understanding data visualization-supported misinformation online. Furthermore, we relate the existing research on misinformation interventions to the problem of visual misinformation.

2.1 Visual Misinformation Online

Prior work has documented the potential of data visualizations to mislead their audience, both through deceptive features of visualization design that interfere with viewers' ability to accurately read off values from a chart [7, 29, 36] and through logical fallacies and confirmation bias that result in visualizations supporting misinformation arguments [23, 28]. Lee et al. [23] discuss that in online COVID-19 discourse, oftentimes pro- and antimask communities have used the same visualizations to argue for opposing views. The multipurpose nature of COVID-19 charts supports the idea that the misleadingness is often not an objective attribute of a visualization, but rather is viewer-dependent. Differences in how viewers interpret the same data visualizations are likely to occur due to a variety of factors, including the social context a viewer is exposed to [15], individual differences [54], and personal biases [38] between viewers, as well as the *curse of knowledge*—an assumption that others

[1]Known as X since July 2023.

interpret the chart the same way you do [50]. Existing research primarily focuses on people's direct reactions to visualizations. Yet, charts shared online typically do not exist in a vacuum but rather are embedded in a post and can be part of a conversation or be accompanied by an interpretation. Furthermore, although any biased framing is known to influence a viewer's reading of a chart [19], a visualization post's text may serve as the main source of misinformation [28]. Therefore, in studying data visualization-supported misinformation, it is important to focus not just on reactions to the visualization itself but also on the (potentially misleading) insight it serves to support. To capture the variety of responses elicited by data visualization interpretations of others, our paper analyzes engagement with data visualization posts on social media, and describes factors that lead online audiences to agree or disagree and trust or distrust such interpretations.

Similar trust factors have been described for textual misinformation before [8, 53], but although work examining the fact that "people lie with charts" goes back decades [14], there has been a dearth of research conceptualizing data visualization-supported misinformation and studying it at the same level as textual misinformation. Recently, however, researchers have started to examine the role data visualizations play in the creation and spread of misinformation, and, importantly, how data visualization-supported misinformation fits in the broader existing research on online information integrity. Weikmann and Lecheler [48] argue that visual disinformation is "its own type of falsehood [that] differs from textual disinformation." The authors discuss that misleading visuals have both higher modal richness than text and are associated with a higher level of manipulative sophistication, resulting in more credible and convincing disinformation [48]. Matthew Hannah, in presenting a case study of QAnon conspiracies online, argues that QAnon's success—and even existence—relies exclusively on the effectiveness of their information visualizations and search for patterns in data [12]. Hannah discusses that this success is "symptomatic of our inability to combat misinformation that mimics the methods of data analysis" [12]. Our work attempts to fill the research gap in understanding "misinformation that mimics the methods of data analysis" by describing the ways the crowd reviews misleading data interpretations.

2.2 Online Misinformation Interventions

With the spread of online misinformation, researchers and social media platforms have been preoccupied with finding ways to design scalable interventions to address the spread of misleading and harmful content. Aghajari et al. [1] present a literature review of existing interventions, categorizing them as content-, source-, user-, and community-oriented. By far the most commonly known type of intervention is content-based, which focuses on the veracity or credibility of the content. Content-based approaches have been implemented by most major social media platforms such as Facebook and Twitter, and include removing, deprioritizing, or labeling content based on its veracity, as determined by expert fact-checkers or an algorithm [1]. As our approach in this study focuses on reviewing the content of posts, we primarily discuss the potential interventions against data-driven misinformation in this paper through the lens of content-based approaches.

Research on the efficacy of fact-checking interventions, however promising, has so far presented heterogeneous results [46]. Even though interventions are often successful in their goal of correcting people's beliefs, researchers have described the potential for fact-checking to have a *backfire effect*: to solidify incorrect beliefs [41] and to increase toxicity [33]. Similarly, the *implied truth effect* may lead the audience to believe that all other, not-yet-fact-checked content is accurate [39]. Crowdsourced fact-checking interventions are a promising way of efficiently scaling up fact-checking [2]. Yet, these interventions come with pitfalls, such as the observation that politically aligned users are unlikely to fact-check each other [3].

The heterogeneity in intervention efficacy research may stem from the fact that the underlying misinformation presents a wide variety of types of misleading statements that we are yet to fully understand and, importantly, distinguish between [47]. Specifically, in their empirical study of fact-checking effectiveness in political news articles, Walter and Salovich [47] find that audiences also struggle to distinguish between opinion- and fact-based pieces, which has a major influence on the effect of misinformation corrections. As many works that design and propose misinformation interventions for social media discuss [16, 17], people especially struggle to correctly assess the "gray area" of misleading but factually accurate statements, such as opinions, incorrect interpretations of data, or satire. Data- and data visualization-driven misinformation is based on factual data with a potentially opinionated interpretation. Studying these forms of misinformation presents an opportunity to fill the research gap in our understanding of engagement with factual but misleading content. In this paper, we argue that *data-driven misinformation* is a distinct type of misinformation that requires special consideration in intervention design.

3 STUDY 1: QUANTITATIVE ANALYSIS OF ENGAGEMENT

To address the question of whether accompanying a data visualization post with an insight—and, moreover, a misleading one—has an effect on audience's engagement with the post (**RQ1**), we conducted a quantitative analysis of engagement. Specifically, this analysis allows us to identify whether misleading data visualizations are associated with being discussed, shared, or liked more than other posts. In this section we describe our approach to data collection and regression analysis and summarize the results of our Study 1.

3.1 Methods

In order to quantitatively analyze the effects of visualization insights on engagement, we used our data to estimate regression model coefficients. In this section, we describe our approach in detail, from engagement data collection to considerations in model selection.

3.1.1 *Data Collection and Processing*

As the basis for our data collection, we used the publicly available data set and supplemental materials from Lisnic et al.'s study of misleading data visualizations on Twitter, which spans the time period between May 15, 2020 and September 6, 2021 [28]. In their data set, the authors provide tweet IDs and the corresponding descriptive variables, such as tweet polarity, presence of reasoning errors, or violations of visualization design guidelines. Of the 9,958 tweets from Lisnic et al. [28], 1,060 have been removed from the

platform or made private by the authors, resulting in 8,898 original tweets used in our analysis.

In order to analyze engagement, we used Twitter API's full-archive search to collect the complete engagement data associated with the original tweets: we collected 668,173 retweets, 229,764 replies, and 101,705 quote tweets for a total of 999,642 *engagement tweets*. To control for tweet author effects in our regression analyses, we additionally collected user data for all tweet authors in our data set to use as covariates, including follower count and verified (or "blue check") status. Our data collection occurred between February and March of 2023, and as such was not affected by the changes to Twitter's verification program from April 2023.

We minimally processed the data by merging engagement tweets and author data with the original tweet data. We provide our data processing scripts as well as tweet IDs of posts used in our analysis in the supplemental materials. To comply with Twitter's API policies, we are unable to provide full tweet data but it may be rehydrated using the IDs, as long as the tweet is still publicly accessible.

3.1.2 *Regression Analysis*

To **analyze the effects of providing accurate or erroneous insights in a data visualization post**, we conducted a regression analysis of count and duration of the main forms of engagement: replies, retweets, quotes, and likes. As our explanatory variables, we used the *opinion* and *reasoning error* data from Lisnic et al. [28]. In our analysis, we use the term *insight* to refer to Lisnic et al.'s opinion variable, which denotes tweets in which the author explicitly highlights or hints at observations, trends, or hypotheses in the data. Non-insight posts share data visualizations without interpretation, such as neutral status updates. Most insights are explicitly stated in the tweet text or added annotations, but some are inferred by holistically analyzing the tweet author's feed and follow-up replies [28].

To model the engagement count variables—the number of replies, quotes, retweets, and likes of a post—we fit Negative Binomial regression models. Negative Binomial regressions are a generalization of Poisson regressions that are commonly used to model count data. Negative Binomial models loosen the assumption of variance being equal to the mean used in Poisson models, and are thus more appropriate for our highly dispersed data, confirmed by the over-dispersion coefficient θ being highly statistically significant in our Negative Binomial regressions. Additionally, we confirmed that Negative Binomial regressions outperformed Poisson on our data by various other model selection criteria, such as Akaike's Information Criteria (AIC), Bayesian Information Criteria (BIC), and Mean Absolute Error (MAE). As a robustness check, we provide the results of Poisson regressions and model selection tests in the supplemental materials, as well as the scripts used to generate them.

Social media engagement data generally tends to be highly right-skewed—with most posts receiving little to no engagement and few posts going viral [10, 21]—which is also the case with our data. One of the sources of high skewness we observe is the high number of zeros in the distribution of the reply counts, with 43% of tweets in our data set having no replies. It is possible that a post may receive zero replies via two mechanisms: structural zeros in posts that signify lack of interest in commenting on a post (or, being the first to comment on a particular post), and random zeros that stem from the fact that the post was not seen by enough people. To account for the excess zeros and model the two ways of generating such excess zeros in our reply data, we fit a Zero-Inflated Negative Binomial (ZINB) regression. A ZINB regression is a type of zero-augmented approach that models a mixture of two distributions: a logistic regression that models generation of zeros, and a Negative Binomial regression that estimates reply count. Zero-inflated regressions are a commonly used way to model social media engagement data [24, 26, 43].

Model	df	LL	AIC	BIC	MAE
Replies					
Zero-Inflated NB	35	**-25,686**	**51,441**	**51,689**	**32.84**
NB	18	-25,908	51,852	51,980	33.32
Retweets					
Zero-Inflated NB	35	**-39,004**	**78,078**	78,327	98.13
NB	18	-39,078	78,193	**78,321**	**98.12**
Quotes					
Zero-Inflated NB	35	**-23,527**	47,124	47,372	**14.93**
NB	18	**-23,527**	**47,090**	**47,218**	**14.93**
Likes					
Zero-Inflated NB	35	**-127,049**	**254,168**	254,447	481.00
NB	18	-127,111	254,258	**254,402**	**480.87**

Table 1: A summary of metrics used to evaluate and compare engagement count model specifications. We compared the fit of Zero-Inflated Negative Binomial (ZINB) and that of regular Negative Binomial (NB) using log-likelihood (LL), Akaike's Information Criterion (AIC), Bayesian Information Criterion (BIC), and Mean Absolute Error (MAE) metrics. We highlight the most accurate performing model by each criterion (lower is better). As seen from the table, the metrics suggest that the Zero-Inflated version of the model provides a better fit for replies, but we observe mixed results for other metrics.

Despite doubling the model complexity, as seen in Table 1, in our model selection tests the ZINB model for reply counts also outperformed the non-zero-augmented approach using Akaike's Information Criteria (AIC) and Bayesian Information Criteria (BIC), which account for the additional model complexity of a ZINB. Table 1 also shows that for other metrics—retweets, quotes, and likes—the zero-inflated approach shows improvement in some metrics but not others, which is expected since their distributions, albeit still skewed and having excess zeros, contain fewer zeros than the replies. For consistency, we present the results of ZINB models for retweets, quotes, and likes as well; however, we note that the coefficients of corresponding non-zero-inflated models are similar and can be found in the supplementary materials.

In addition to engagement counts, we also investigated **the effect of data insights on the duration of the post's engagement**. Duration of engagement is calculated as time elapsed in hours between the original post and the latest reply, retweet, or quote tweet as of February 2023. Since the Twitter API does not provide timestamps of individual like events, we are unable to make inferences about duration of likes for posts. To model engagement duration (a continuous variable rather than a count variable), we fit standard

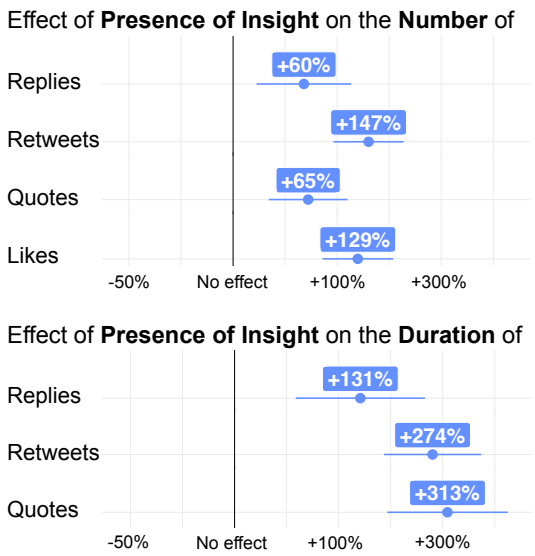

Figure 2: Average effect of presence of insight (compared to no insight in a post). Regression coefficients were estimated controlling for presence of reasoning error, effects of which are presented in Figure 3. We show 95% confidence intervals of estimated effect size of variable on count and duration. Estimated effects are calculated as $e^{\beta}-1$, where β is the output regression coefficient. We observe that the presence of an insight in a post is associated with a higher number and longer duration of engagement.

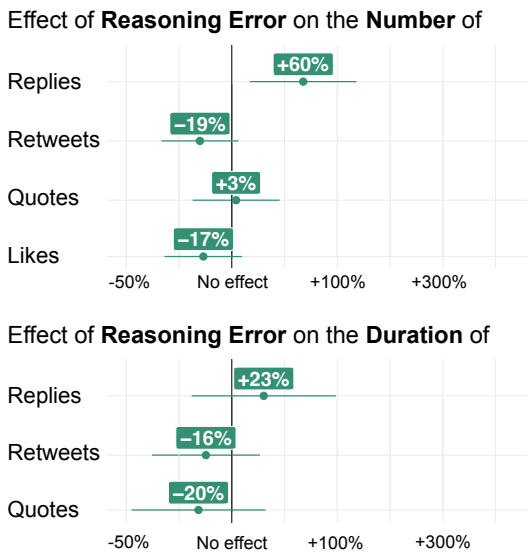

Figure 3: Average effect of presence of reasoning error in insight (compared to insight with no reasoning error). Regression coefficients were estimated controlling for presence of insight, effects of which are presented in Figure 2. We show 95% confidence intervals of estimated effect size of variable on count and duration. Estimated effects are calculated as $e^{\beta}-1$, where β is the output regression coefficient.

multiple linear regression models with log-transformed response variable, to account for the skewness.

The results of regressions presented in this paper correspond to models that control for author-, visualization-, text-, and time-specific covariates. *Author features* include (log-transformed) number of followers and verified status. *Visualization features* describe whether the attached data visualization is a screenshot of an existing chart, has any author-added annotation, or has any violations of common visualization design guidelines (e.g., truncated, inverted, or dual axes). *Text features* control for the number of words in the tweet, as well as the number of mentions, emojis, hashtags, and external URLs. *Time features* include weekend and time-of-day fixed effects, separated into four six-hour segments. In the interest of robustness, we calculated the results excluding different sets of covariates and note that the statistical significance and magnitude of observed effects are consistent across model runs.

3.2 Results

Figures 2 and 3 show results of the Negative Binomial regressions of engagement counts as well as the logged duration of engagement regressions, respectively.

3.2.1 *Engagement Count*

Based on the results shown in Figure 2, we observe that **data visualization posts that provide an insight** by offering an interpretation or pointing out a specific aspect of the chart (as opposed to simply sharing a chart) **are associated with significantly higher levels of all forms of engagement.** Specifically, our results show that providing an insight is associated with, on average, 60% more replies, 147% more retweets, 65% more quotes, and 129% more likes.

As seen in Figure 3, an erroneous insight in a post is associated with an additional 60% more replies. The effect of errors on other types of engagement—such as retweets or likes—is limited in effect size and statistical significance. In other words, **an erroneous data interpretation attracts significantly *more discussion* while having *no effect on the breadth of spread.***

3.2.2 *Engagement Duration*

We find that providing an insight is associated with longer-lasting engagement and conversations: as seen from Figure 2 our model with the complete set of covariates shows that posts with insights are associated with, on average, 131% longer duration of replies, 274% longer duration of retweets, and 313% longer duration of quotes.

The results of our duration regressions in Figure 3 also indicate that there is not a similar effect of reasoning errors on the longevity of engagement. We find, on average, slightly longer duration in replies and slightly shorter in retweets and quotes; however, the effect sizes and levels of significance are relatively low.

In summary, our results show that data visualization posts with insight remain relevant for a much longer time than those without. In the context of COVID-19, we speculate, based on these findings, that visualizations without an insight are used as status updates and provide the latest statistics that may be relevant for only one day (median of 14 hours). At the same time, posts with interpretations use the same data to tackle more fundamental questions, garnering discussions that last multiple days (median of 29 hours).

4 STUDY 2: THEMATIC ANALYSIS OF ENGAGEMENT

Following the results of Study 1, we set out to explore the contents of replies to posts with erroneous insights and investigate whether online audiences are able to identify and raise awareness about misleading data interpretations as evidenced by the content of their replies (**RQ2**). To do so, we performed a thematic analysis of direct engagement in a subset of our data. In this section we describe our approach and summarize the results of Study 2.

4.1 Methods

With the goal of qualitatively analyzing the engagement with data visualization insights, we performed *template analysis* [18] to construct a hierarchical code book that describes the content of replies and quotes of posts in our data. This section outlines our process in detail, from selecting a sample of data for thematic analysis to performing quality and reflexivity checks.

4.1.1 *Data Selection*

To select a sample of data that is large enough to allow us to identify important themes yet small enough to be analyzed it in depth, we performed multistage stratified sampling. Firstly, we filtered our data set to posts that contain an *insight*—observations, trends, or hypotheses in the data highlighted by the author [4]. In the original data set, Lisnic et al. [28] use the term *opinion tweets* for this concept. These are the posts that are, by definition, amenable to being misleading and therefore are the focus of our engagement analysis.

Secondly, to limit our data to relevant engagement with the original post in question, we selected all first-level reply posts or quote posts, except for those authored by the same user as the original post. These posts form a set of all posts that *directly engage* with the original post, as opposed to replies or author's own follow-ups or threads. Thirdly, with the goal of reviewing a richer variety of responses, we excluded posts with fewer than 16 direct engagements, which is the median value among posts with any direct engagement. Lastly, to reduce our sample for thematic analysis, we randomly sampled 30 posts with a reasoning error and 30 without, for a total of 60 original posts with median-or-above engagement count. We then used all of their associated 3,806 first-level replies or quotes for our thematic analysis.

4.1.2 *Template Analysis*

Our approach to thematic analysis was guided by the template analysis techniques described by King [18]. In choosing a methodology for our thematic analysis of engagement with data visualizations, our goal was to strike a balance between the structure of "small q" qualitative methods that emphasize development of coding schemes, and a more contextual and reflexive analysis of themes offered by "Big Q" qualitative approaches, as described by Braun and Clarke [5]. In the context of this research, we wanted to acknowledge the participatory role of the researcher and our research goals, as well as our interpretation of the cultural and semantic context of social media discourse in our conceptualization of themes, while leveraging a structured code book to assist us in describing individual tweets—a relatively independent and small unit of analysis. At its core, template analysis involves developing a code book called a *template* in a way similar to more positivist and postpositivist approaches; however, the template is used as a tool to help the researcher scaffold data and conceptualize themes rather than a way to convert qualitative into quantitative data [5, 6, 18].

The process of developing the coding template started with the first author reviewing a random sample of 500 first-level replies and noting an initial set of codes. Although we generated most of our code book inductively, in order to more efficiently process our large data set, we deductively defined a set of a priori codes [18] based on existing literature and our own domain knowledge. The lens through which we developed the initial set of codes was guided by the authors' interest in examining how social media audiences review or fact-check misleading data visualization posts. Consequently, our thematic analysis is influenced by the initial code book's direction and pays special attention to users' general analytical engagement with data and data insights, rather than specifics particular to the topic of the posts, COVID-19 data. In the next step, the first author reviewed the complete set of 3,806 direct engagement posts, iteratively revising the contents and structure of the code book. Lastly, the authors used subsets of the code book to conceptualize themes by highlighting and contrasting higher order categories of codes from the final template.

With the goal of validating and scrutinizing the analysis, we performed two iterations of quality and reflexivity checks, as described by King [18]. The first check occurred after development of an initial template and involved a coder independently coding 500 randomly selected posts using the initial template. The first author met with the coder to discuss whether the codes were straightforward to apply, whether the data was easily described by the codes, and whether the template failed to capture any relevant themes. As a result, a new theme related to audience's communication of trust was conceptualized and the template was adjusted for clarity. The second check occurred after the first author completed reviewing the full data set and developed an updated template. In the second check, two senior authors independently coded different subsets of 100 posts each. All the authors met twice, once in the middle of the check and once at the end, to discuss the clarity and richness of the template. Following the second quality check, no new themes were conceptualized, yet several template items were updated in name and definition to more broadly describe the data.

After conducting the second quality check, the authors agreed that the template provides a sufficiently good and rich representation of the themes we identified in the data. The final coding template is presented in Figure 4. We provide an audit trail of the evolution of our template in the supplementary materials. The themes presented below were synthesized through interpreting the final template, noting insightful differences and similarities between individual codes or sets of codes.

4.2 Themes

In this section we present the results of our thematic analysis. For each theme we describe how it relates to specific codes or groups of codes from the template in Figure 4 and illustrate it with examples from our data. The examples of posts and replies presented throughout the paper are minimally edited to fix typos and remove usernames to preserve anonymity. We then offer a discussion of

1. **Sentiment**
   - **1.1. (Dis)trust in insight**
     - 1.1.1. Explicitly or implicitly agreeing
     - 1.1.2. Suggesting a conspiracy
     - 1.1.3. (Dis)trust of data/source
     - 1.1.4. (Dis)trust of statistics/visualization
     - 1.1.5. Appeal to facts
     - 1.1.6. Sharing (by quoting or tagging)
     - 1.1.7. Meme/joke
     - 1.1.8. Mocking caveats
   - **1.2. (Dis)trust in poster**
     - 1.2.1. Asking for advice/elaboration
     - 1.2.2. Asking for more/updated data
     - 1.2.3. Asking for source
     - 1.2.4. Gratitude/respect
     - 1.2.5. Lack/presence of expertise
     - 1.2.6. Personal attacks
   - **1.3. Direction**
     - 1.3.1. Trust
     - 1.3.2. Distrust
2. **Content**
   - **2.1. (Quasi)analytical**
     - 2.1.1. Anecdote
     - 2.1.2. More data
     - 2.1.3. Caveat
     - 2.1.4. Reinterpretation
     - 2.1.5. Update
     - 2.1.6. General caution
   - **2.2. Citations**
     - 2.2.1. None
     - 2.2.2. Visualization
     - 2.2.3. Raw data
     - 2.2.4. Article
     - 2.2.5. Authority figure
   - **2.3. Attempts to fact-check**
     - 2.3.1. Revisiting
     - 2.3.2. Fact-checking the non-data part
     - 2.3.3. Redirect to authority figure
   - **2.4. Direction**
     - 2.4.1. Uphold/strengthen insight
     - 2.4.2. Oppose/weaken insight

Figure 4: Final template used to describe the data and conceptualize themes. The codes under 1. Sentiment describe users' trust in the poster or in the general sentiment of the post. The codes under 2. Content describe the replies' analytical engagement with the data and visualization.

the implications of the relevant findings of the theme in the context of designing interventions against data-driven misinformation. To conclude, we summarize our discussion by identifying the **opportunities** that the theme presents to effectively address misinformation and describing important **limitations** of the opportunity.

4.2.1 *Analytical Wisdom of the Crowds*

Based on our thematic analysis, we identify evidence that online crowds can and do reason about the accuracy or misleadingness of data visualization posts and analytically engage with the data and its interpretation. As seen from the subitems in code 2.1 in the final template in Figure 4, we observe six ways in which the audiences analytically assess the data interpretations in their response: sharing personal anecdotes or lived experiences that add context to the data (2.1.1. Anecdote), providing more data points of the same metric or a different variable (2.1.2. More data), highlighting important statistical or methodological caveats (2.1.3. Caveat), reinterpreting the original chart to underscore a different insight (2.1.4. Reinterpretation), raising awareness about the existence of more up-to-date and sufficiently different version of the data or the chart (2.1.5. Update), and generally cautioning against making strong conclusions based on limited data (2.1.6. General caution).

Notably, such analytical assessments not only serve to undermine and "fact-check" the original insight but also can be used to confirm or strengthen it, indicated by Direction codes 2.4.1. (uphold) and 2.4.2. (oppose). For instance, sharing a larger set of data points may highlight the fact that the original visualization was cherry-picked if the trend is different, or it could provide evidence that it was not if the trend is consistent. Similarly, sharing a methodological caveat, e.g., that the recording of COVID-19 cases is delayed and thus undercounted, can weaken an insight that highlights a dip in cases but further strengthen an insight that highlights an increase.

Implications. Our findings describe the avenues that a crowd of nonexperts has to analytically assess the accuracy of a data interpretation on social media. In our data set we do not observe users sharing specialized domain knowledge stemming from their expertise, performing original research, or surveying existing research—which is to be expected of a majority nonexpert crowd in a fast-paced microblogging environment. Instead, users rely on their own lived experience and individual pieces of information or data already familiar to them to interpret or reinterpret the original conclusion. As a result users are likely biased by the information readily available to them.

A significant limitation is that individual lived experiences or counter-data cannot entirely disprove the original insight. Moreover, the crowd's assessments also cannot accurately estimate the extent to which a given caveat impacts the insight. For instance, the caveat that the vaccine adverse effects system (a web-platform to track adverse effects) allows unverified submission from anyone in Figure 7 suggests that cases of vaccine-related deaths and adverse effects are likely overcounted. However, since this caveat is merely directional and does not provide any information about by *how much* the cases are overcounted, we cannot know if the original insight still holds. Effectively, the audience's analytical assessments can be fruitful in sowing doubt and undermining trust in the original conclusion but cannot disprove it.

Opportunities: Non-expert online audiences identify important and nuanced caveats in misleading data interpretations.

Limitations: Caveats cannot fully disprove flawed data interpretations, only weaken them or sow doubt.

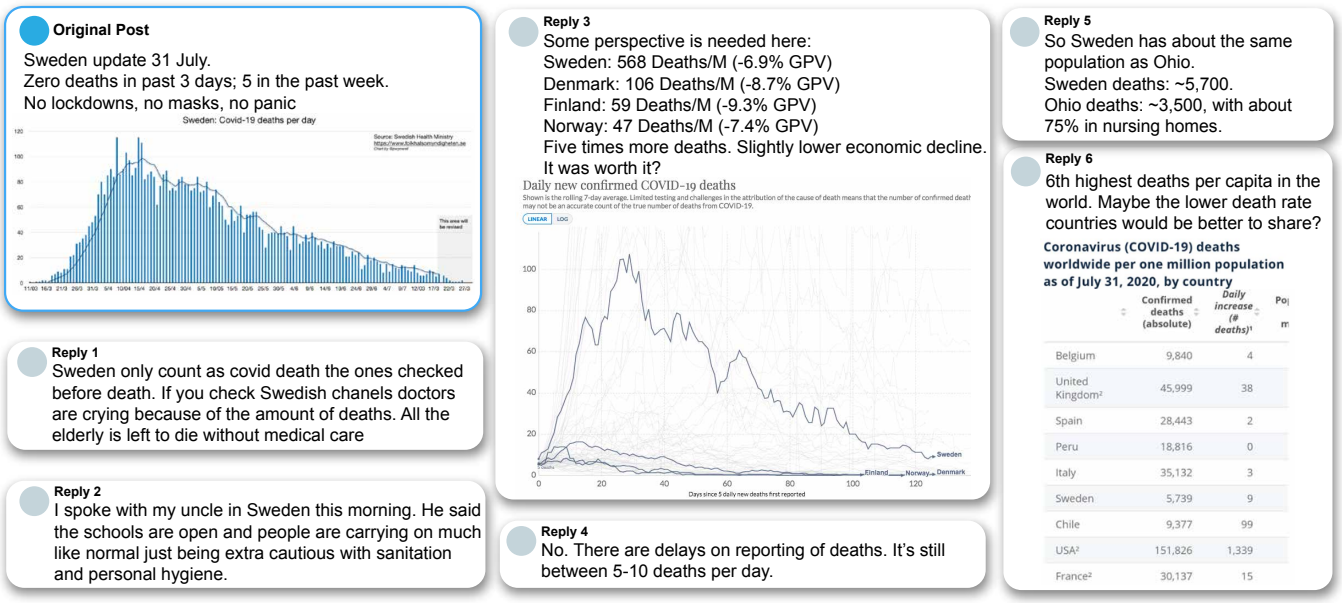

Figure 5: Example post where the author promotes the idea that COVID-19 containment measures, such as masking, are ineffective citing the data that shows death per day going down in Sweden. The replies to the post showcase the types of analytical responses from the crowd that challenge the accuracy and generalizability of the author's conclusion: sharing of more data, caveats, up-to-date data, and personal anecdotes.

4.2.2 *Debunking Is in the Eye of the Beholder*

We identified an important difference between an audience agreeing with the *premise of the post* and agreeing with the *presented analysis or data interpretation*. Consequently, users are able to find fault with the particulars of the data while still upholding the conclusion, with one reply stating: *"Yeah, this graph doesn't show that, but we get the point."* In the code book this difference is highlighted by groups of Codes 1.3. and 2.4. seen in Figure 4: codes in 1.3. describe the direction of trust, or whether the reply trusts the author's expertise and insight, whereas codes in 2.4. describe whether any analytical assessment strengthens or weakens this insight.

In another example, the audience proactively seeks to build on a flawed analysis they agree with by suggesting improvements: the post in Figure 6 attempts to highlight the effectiveness of vaccines against COVID-19 by sharing statistics of cases during a local outbreak. Numerous responses call attention to the fact that the interpretation is flawed due to base rate fallacy—the author did not share population-level statistics, only those pertaining to existing patients. Yet at the same time, most replies find it important to note that although they are pointing out this fallacy, they are in full support of vaccination and agree with the author's conclusion. One reply notes, *"I've been vaccinated. Just not one for misleading data."* At the same time, we observe explicit or implicit hesitation when commenters challenge an insight they agree with. As one respondent puts it, *"I can find holes in this graph but I won't because I want people to wear masks."*

Implications. This finding calls attention to an important difference between assessments of data-driven misinformation and factual statement-based misinformation. Previous work by Allen et al. finds that, in the context of factual statement-based misinformation, politically aligned users are unlikely to formally fact-check each other [3]. However, although a factual statement can be true or false, Lisnic et al. discuss that most misleading data visualization-supported arguments take the form of an inductive argument, which can be plausible or implausible [28]. As a result, it is possible to arrive at a correct conclusion even through a flawed analysis of data, and consequently it is possible to challenge the analysis without debunking the conclusion.

We still, however, observe evidence that like-minded users are sometimes hesitant to probe flawed data interpretations. This observation highlights a limitation in the crowd's ability to effectively evaluate the accuracy of data-driven insights: a large portion of a post's audience may forego their assessment of the *analysis* due to concerns about unintentionally convincing others that the *conclusion* is false. As a result, analytical assessments are mostly submitted by users who disagree with the conclusion and attempt to attack it. Thus, submitting a flawed analysis to support a true conclusion may backfire and do a disservice to the conclusion: most replies are likely to be attacking the insight and inadvertently convincing others that it is wrong altogether.

Opportunities: Users who agree with the conclusion still often point out that the analysis is misleading attempting to strengthen it.

Limitations: Nonetheless, ideologically aligned users appear to be more hesitant to share their assessments.

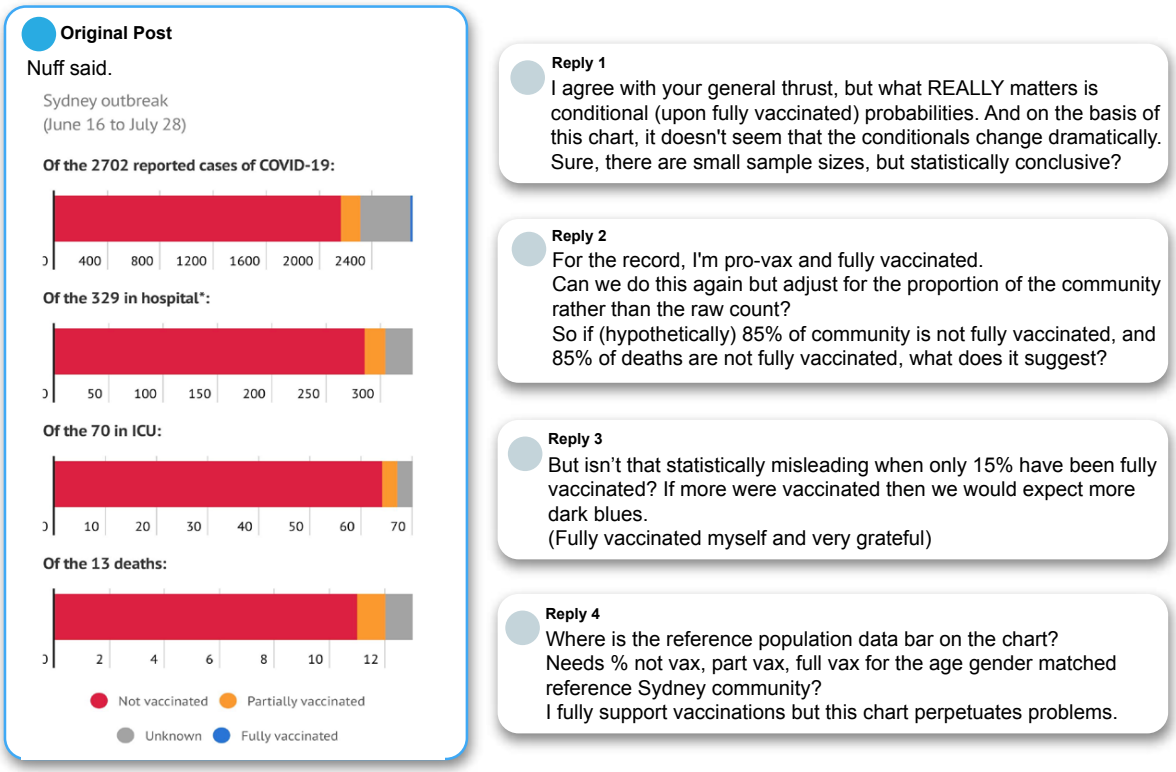

Figure 6: Example post with replies showing the types of analytical responses from the crowd. The responses are predominantly agreeing with the conclusion, yet still point out flaws in the data interpretation.

4.2.3 *What Cannot Be Fact-Checked Could Be Peer-Reviewed*

Misleading data-driven insights leave few opportunities for audiences to share a statement that would, if true, prove the invalidity of the claim—or to "fact-check" it. Although fact-checking is common in cases of fact-based misinformation, visualizations insights typically take the form of data-supported hypotheses. In our analysis we identify limited cases in which audiences attempt to fact-check data-driven insights, listed as Codes 2.3.1. through 2.3.3. in Figure 4. In cases when the visualization is outdated, sharing new data could invalidate the original insight (2.3.1. Revisiting); in cases when the insight is true only with the addition of a nondata statement (for instance, a false claim that the FDA approved the use of a certain drug against COVID-19), that statement itself could be fact-checked (2.3.2. Fact-checking the nondata part); and lastly, some users attempt to invalidate a data-driven insight by sharing repudiating quotes and official statements from people in positions of authority, such as politicians or scientists (2.3.3. Redirect to authority figure).

Predominantly, however, misleading visualization insights in our data cannot be invalidated by a single response. As discussed in Section 4.2.1 and represented by Codes 2.1.1. through 2.1.6. in Figure 4, users attempt to contest misleading insights by sharing a single piece of counter-evidence or a caveat to the claim. In our analysis we note that although an individual user's response only provides *one piece of evidence* that often does not disprove the claim on its own, reviewing the entire conversation reveals a *variety* of independent pieces of evidence that form a consensus. For instance, the post in Figure 5 makes an argument that lockdowns are ineffective because Sweden—a country that did not have a strict lockdown—is experiencing a dip in cases. The responses point out a variety of possible counterarguments: the caveat that Sweden allegedly undercounts deaths, additional data showing that Sweden has more cases than comparable Nordic countries and even than most other countries in the world, the caveat that death counts for recent dates are delayed, or personal anecdotes of locals reporting that they are still "cautious with sanitation and personal hygiene" despite a lack of formal lockdowns. Thus, a viewer is presented with vastly more evidence against the original insight than in support of it.

Implications. In Section 4.2.1 we discussed that it is typically not possible to estimate the extent to which an individual analytical assessment impacts the original insight. Evaluating the whole set of replies, however, may communicate a more complete assessment of the original claim: if multiple unrelated pieces of evidence point out the incompletenesses of the insight, it is likely that the insight is misleading. The process of individual users reviewing the accuracy of the original interpretation is akin to *crowd peer-review* or formation of a *crowd consensus* on the topic. A diverse crowd offers a wide variety of lived experience, domain knowledge, and data and statistical literacy, and contributes what they know best—usually only a single detail—to the conversation. Consequently, no single reply contains a complete assessment of the original post, but the entire conversation serves as the crowd's assessment.

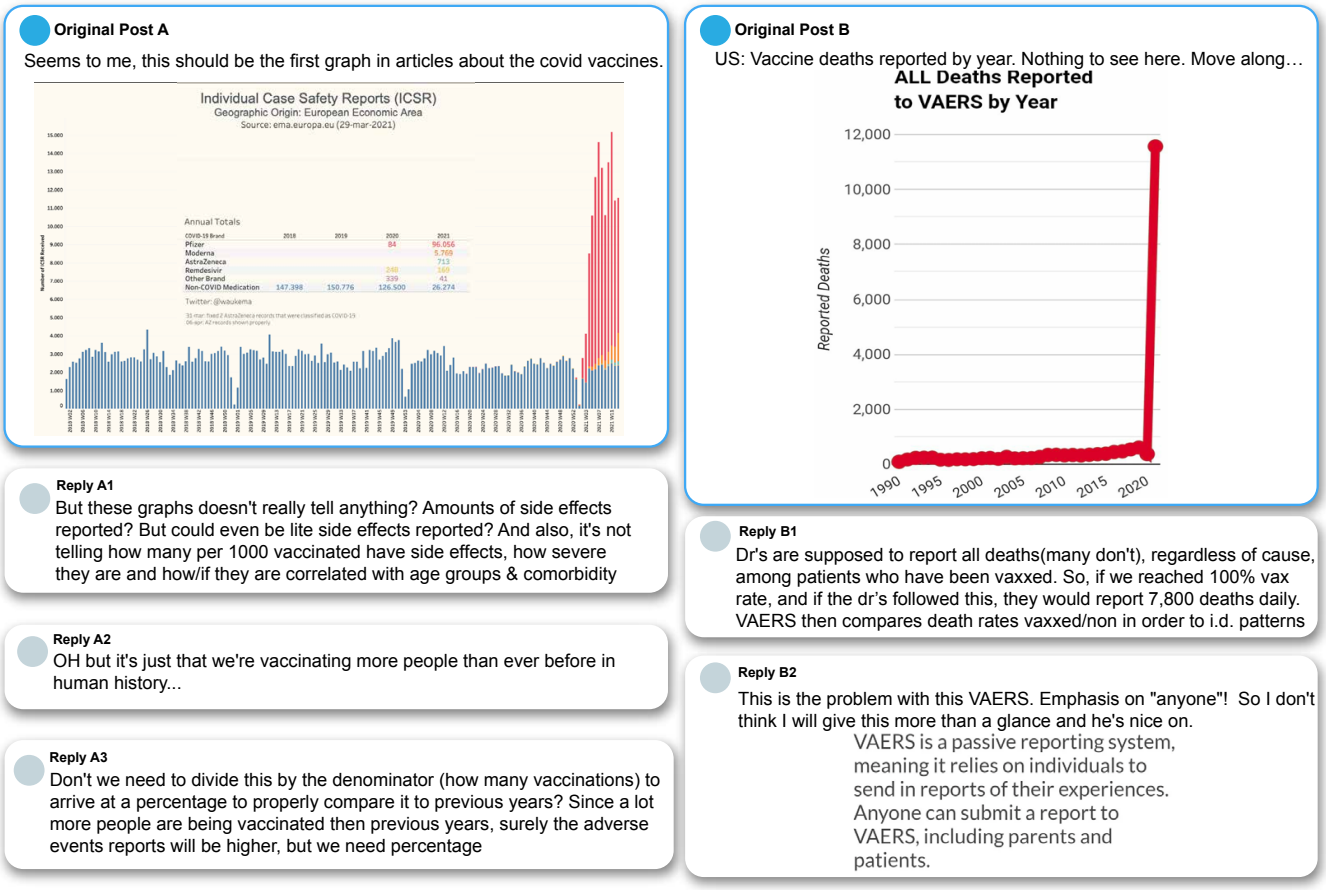

Figure 7: Example posts using similar data—EU's ICSA and US's VAERS adverse effects tracking systems—to spread skepticism around safety of vaccines. These databases have been widely misinterpreted by antivaccine activists to promote their views [45]. The responses point out caveats in the interpretation, such as the need to account for the fact that there are mass vaccinations underway, and data limitations like the lack of concrete definition of "adverse effect" and, most importantly, the fact that the submissions are not verified and can be submitted by anyone.

Although our findings indicate an opportunity to leverage the hive-mind for a crowd peer review of misleading data interpretations, there are challenges. To be used effectively, the assessments from the entire conversation body need to be surveyed and synthesized into a *meta-review* that presents the diverse points of view. It is also necessary for the body of "reviewers" to be large and diverse, which is difficult to achieve for posts that do not go viral or authors with a highly partisan audience.

Opportunities: Longer discussions of posts with misleading data interpretations cover a diverse set of caveats, counterexamples, and anecdotes.

Limitations: To leverage the replies to (in)validate data insight, a large and diverse audience is required; and many individual assessments must be synthesized into a "meta-review" to present a complete picture.

4.2.4 *Data Does Not Speak For Itself*

Up to this point, our highlighted themes have focused on the audience's engagement with the analytical content of data interpretations. However, whereas analytical soundness of a data visualization insight is an important consideration of credibility brought up by the replies, we identify other credibility factors that exist independent of the insight itself. Codes grouped under Items 1.1. and 1.2. in Figure 4 describe a variety of explicit and implicit indications of trust and distrust of the author or insight shared by the replies, including trust or distrust in data integrity or data sources (1.1.3.), perceived level of data literacy or domain expertise of the original author (1.2.5.), or personal attitude about the author unrelated to the analysis (1.2.6.).

Examples in Figure 8 indicate that the lack of a source for the data or chart negatively affects its credibility (as one user noted sarcastically: "No source. Seems legit..."). At the same time, presence of a source a user disagrees with—whether it is "Florida and Texas" or "liberals"—can also lead to an insight being dismissed and distrusted. Furthermore, users often distrust some data visualization

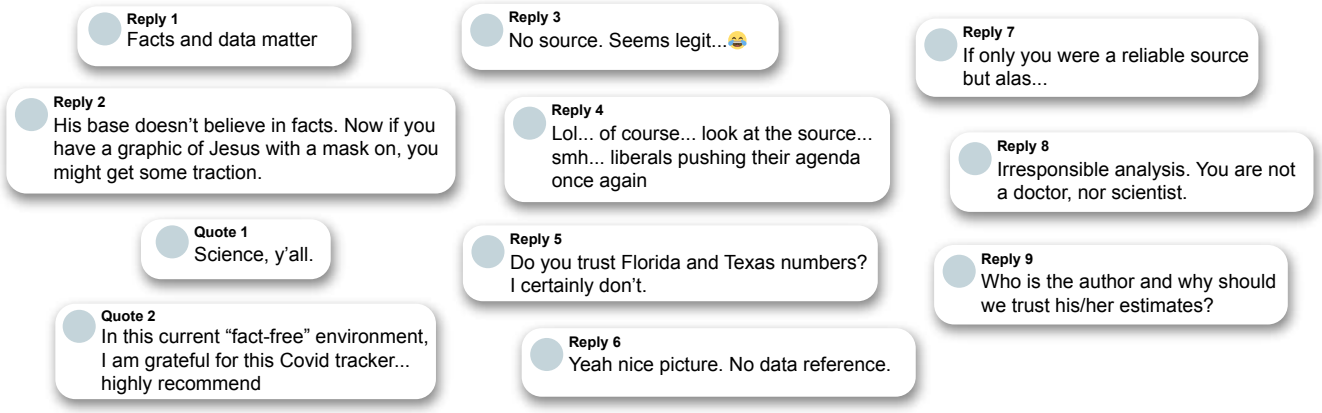

Figure 8: Examples of replies to and quotes about a variety of different posts with data-driven insights that indicate sentiment toward the author or the insight without analytically evaluating the insight. Examples include replies that trust data insights because they are based on "facts", or replies that distrust data insights because of their doubts about source validity or the author's expertise and credentials.

posts because they are aware of the fact that statistics can be presented in a misleading way, whereas others compare claims backed by data to "facts." Such replies do not analytically engage with the chart or the insight itself, pointing to the variety of credibility and trust factors beyond the content of the original post.

Implications. Data or its visual presentations do not exist in a vacuum but rather are entangled with the social media persona sharing it as well as the existing conspiracies and stereotypes concerning the topic of interest. Our results indicate that in many cases users exhibit such a strong sense of trust or distrust of the author or the data source that they do not feel the need to analytically engage with the data insight to decide whether they believe it.

Our findings highlight the flexible nature of using data as evidence of phenomena: although users often advocate for democratizing data, "doing one's own research," and compare data to "facts" (Code 1.1.5.), other examples indicate that being "a doctor [or] a scientist" is an important prerequisite for sharing data-driven insights. This consideration is important for effective scientific or public health communication: beyond sharing timely and accurate insights, there is a need for continuous trust-building and engagement with the audience and transparency in data collection and processing methods.

Opportunities: Users question unreliable sources and biased authors and recognize the potential for visualizations and statistics to mislead even with accurate data.

Limitations: On the other hand, users may blindly accept flawed analyses posted by authors they trust.

5 DISCUSSION AND IMPLICATIONS

In this section we discuss our studies' findings and the implications in the context of existing work on data-driven misinformation and interventions. Based on our findings, we offer potential solutions to the limitations and challenges described in Section 4.2.

5.1 Data-Driven Misinformation

The results of our work underscore important affordances and challenges that *data-driven forms of misinformation* present. Misleading data insights fall under a category of posts that Walter and Salovich describe as a 'gray area' of statements that sound like fact-based claims but are actually unverifiable opinions" [47]. It would be unjust, however, to merely call a data-driven insight an opinion. Basing a claim in data offers a veneer of impartiality and scientific rigor, making it more believable than an opinion. And although a data-driven insight is not nearly as certain as a fact, herein lies yet another factor that makes it easier to spread misinformation: it is typically not completely verifiable.

The issues of confirmation or falsifiability of data-driven insights, such as claims of causal relationships, are of course not unique to conspiracy theories shared online. By and large, most of scientific advancements and policy decisions are based in similarly "useful-but-not-certain" data findings—albeit typically with more rigor, confirmatory experimentation, and, more importantly, an admission of uncertainty about the results. In their essay discussing the epistemology of fact-checking in the context of political science, Uscinski and Butler note that fact-checkers' attempts to assess the veracity of causal claims and predictions are futile because even after thorough research many "scientists would be hesitant to dichotomize [such claims] as true or false" [44].

In the world of scientific research, this ambiguity is typically resolved by the community of researchers reaching a scientific consensus. Before a consensus is reached, researchers merely accumulate what Kuhn describes as a "morass" of random facts and unverified observations in hopes that something will show "significant promise for future problem-solving" [20]. Only when a community

forms a *settled paradigm* can researchers perform "normal science": actually advancing the existing theory as opposed to challenging it [20]. Thus, the process of establishing a consensus is highly social and amounts to, through a period of debates, reaching an agreement that a given theory or opinion reflects a current best guess [22].

Although a best guess definitively does not equate with truth, it is useful to present it as highly certain, if not fact. In her exploration of the scientific consensus around climate change, Naomi Oreskes argues that we should treat opinions that the scientific literature largely agrees upon as facts [34]. Oreskes states that excessively communicating stipulations about the uncertainty of scientific findings—amplified by malicious actors who attempt to exaggerate the level of uncertainty—has resulted in general inaction on a variety of topics, such as anthropogenic climate change and the dangers of smoking [34, 35]. As a result, the public severely underestimates the high level of agreement among scientists on a number of seemingly controversial topics, resulting in policy paralysis: oftentimes, scientific consensus is followed by decades of inaction—and the associated costs of inaction—until an idea becomes publicly accepted as fact [40].

It follows that *data-driven misinformation* is most effective at exactly that: forming an illusion of scientific debate and sowing doubt in the existence of actual scientific consensus on a topic. The results of our quantitative study show evidence that data-driven insights with reasoning errors do incentivize such debates by attracting, on average, 60% more engagement that lasts 23% longer. Although these insights based on logical fallacies and spurious correlations may not always succeed in convincing the audience of their claim and forming a new dominant scientific paradigm, they may be persuasive enough in showing that science is not settled on a given topic.

5.2 Designing Interventions Against Data-Driven Misinformation

Based on the above, we argue that, in designing interventions against data-driven misinformation, platforms should be especially cognizant of considerations about data-driven insights being presented as and treated as facts or opinions. In his article discussing the role of facts in modern data-driven discourse, Sun-ha Hong [13] argues that the term *fact* is being overused and mythologized. Specifically, Hong identifies two common practices: *fact signaling*, or performative invocations of facts to discredit rivals and create an "evidence theatre" with data as props, and *fact nostalgia*, an imagined past when "facts were facts." Taken together, these two practices are commonly utilized by actors who spread misinformation to not only present data that support their arguments as facts but also through this process to evoke nostalgia for a mythologized past in which the society had a mutual understanding of what is true and what is false. Consequently, presenting caveats to data-driven insights as fact-checking may have the unintended effect of perpetuating fact signaling and endorsing a dichotomized world that lacks nuance and in which data is either true or false. Uscinski and Butler [44] similarly argue that "[fact-checking] practices share the tacit presupposition that there cannot be genuine political debate about facts, because facts are unambiguous and not subject to interpretation." By being a partial and imperfect representation of phenomena [27], data is often inherently ambiguous and requires contextual knowledge for an accurate interpretation. Hence, instead of presenting a rebuttal as fact, interventions against data-driven misinformation should communicate the ambiguous nature of data by highlighting the limitations of data-driven reasoning and the considerations in attempting to model complex real-world phenomena.

At the same time, if we avoid appealing to facts, we should be careful to not uphold the illusion of the existence of debate and lack of scientific consensus on many scientifically settled controversial topics, such as anthropogenic climate change and vaccine safety. This is a difficult balancing act that involves making a decision about which topics have or do not have scientific consensus. Ways of determining (and proving) the existence of consensus can range from examining literature surveys, consensus conferences [11], and publications such as Cochrane Reviews [25] to data-driven approaches that quantitatively estimate convergence in a network of scholarly literature [40]. We note that in our study we did not observe users attempting to appeal to scientific consensus. This finding could be, to an extent, influenced by the fact that COVID-19 is a novel virus, many aspects of which were, and still are, scientifically inconclusive. To our knowledge, however, existing credibility assessment interventions on social media platforms do not offer a way to raise awareness about scientific consensus, and instead confine the user to a dichotomy of factual correctness that may be confusing in this context. We argue that **the option to appeal to and cite scientific consensus** should be a salient suggestion in the platform's misinformation reporting interface and not make a user decide whether, for instance, anthropogenic climate change is a fact or an opinion.

Our study shows evidence that online crowds do actively attempt to correct data-driven misinformation and are most effective at identifying and highlighting nuances and counter-examples to data insights. We argue that interventions against data-driven misinformation should leverage the strengths of the crowd, and to do so effectively they should address the limitations we outlined in Section 4.2. Specifically, to account for the fact that an individual caveat outlined in a reply is not sufficient to disprove a claim, platforms **should support the creation of meta-reviews of data insights** that summarize the multitude of nuances described by the entire audience. These reviews could be compiled manually by a moderator, by leveraging natural language processing techniques, or through interventions that assist collaborative judgements [9]. Additionally, platforms should **encourage users to share their suggestions for improvements of data interpretations they agree with** to counteract the potential of a backfire effect of flawed analyses in support of true claims. Platforms should also encourage users to share counter-analyses of data as a way of correcting misleading insights by **showing that the opposite conclusion is more strongly-supported**, and go beyond simply pointing out inconsistencies of the original insight.

Besides incentivizing "good data work" and disincentivizing "bad data work," we acknowledge the existence of important credibility indicators of data-driven insights that go beyond the accuracy of the analysis. Based on our findings, we argue that content creators—especially government- and domain-expert-run accounts—should actively work to **build trust in their data and presentation** by being transparent about data sources and collection methodologies and forthright about important data processing decisions. Since conversations surrounding posts with data-driven insights last more

than twice as long as those for other visualization posts, expert accounts should communicate these details by continuously **engaging with the community** and directly addressing concerns raised about the trustworthiness of their insights.

In summary, our overarching recommendation for designing interventions is recognizing data-driven misinformation as a unique and nuanced threat to the integrity of our information space. Misleading data-driven insights undermine the public's trust in scientific findings and promote harmful misinformation while—by the virtue of straddling the line between facts and opinions—remaining largely unaddressed. Through raising awareness about the nuanced spectrum of weak and strong evidence of phenomena, we can tackle the issue of false dichotomies that a claim can only be either fact or opinion or either true or false.

6 LIMITATIONS

Our work is subject to several limitations. Firstly, our data set consisted of content from one platform—Twitter—and thus our findings are influenced by the platform affordances. For instance, character length limits of posts and replies have the potential to limit the amount of detail users share in a single tweet. Additionally, Twitter does not have a variety of features common in message board-type social media sites that could be used to moderate caveats to data-driven insights, such as mega threads or reply pinning. Secondly, our analysis is limited to posts related to the COVID-19 pandemic. Although the initial outbreak of COVID-19 generated a large amount of rich data-driven discussions online, it is also a unique event that featured a lack of existing research on the topic and a high level of politicization. We believe that although such events happen rarely, studying the ways to mitigate the spread of misinformation during them is of utmost importance.

7 CONCLUSION AND FUTURE WORK

In this paper, we presented an analysis of the count, duration, and content of engagement with misleading data visualizations on social media. We hope our work inspires future research to formally study the distinct ways in which data-driven misinformation is generated, spread, and, we hope, corrected. Future work should investigate the impacts of platform affordances on the data-driven discourse by considering other social media sites, as well as the opportunities to address misinformation on various other data-driven topics, such as anthropogenic climate change and vaccine hesitancy. Additionally, future research should identify relevant factors that foster analytical assessments of data-driven insights in a post's discussion beyond the presence of a large and diverse audience.

ACKNOWLEDGMENTS

We wish to thank Cole Polychronis for his contributions. This work is supported by the National Science Foundation (IIS 2041136).

REFERENCES

[1] Zhila Aghajari, Eric P. S. Baumer, and Dominic DiFranzo. 2023. Reviewing Interventions to Address Misinformation: The Need to Expand Our Vision Beyond an Individualistic Focus. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–34. https://doi.org/10.1145/3579520

[2] Jennifer Allen, Antonio A. Arechar, Gordon Pennycook, and David G. Rand. 2021. Scaling up Fact-Checking Using the Wisdom of Crowds. *Science Advances* 7, 36 (2021), eabf4393. https://doi.org/10.1126/sciadv.abf4393

[3] Jennifer Allen, Cameron Martel, and David G Rand. 2022. Birds of a Feather Don't Fact-Check Each Other: Partisanship and the Evaluation of News in Twitter's Birdwatch Crowdsourced Fact-Checking Program. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–19. https://doi.org/10.1145/3491102.3502040

[4] Leilani Battle and Alvitta Ottley. 2023. What Exactly Is an Insight? A Literature Review. In *2023 IEEE Visualization and Visual Analytics (VIS)*. IEEE, Melbourne, Australia, 91–95. https://doi.org/10.1109/VIS54172.2023.00027

[5] Virginia Braun and Victoria Clarke. 2022. *Thematic Analysis: A Practical Guide*. SAGE, Los Angeles.

[6] Virginia Braun, Victoria Clarke, Nikki Hayfield, and Gareth Terry. 2019. Thematic Analysis. In *Handbook of Research Methods in Health Social Sciences*, Pranee Liamputtong (Ed.). Springer Singapore, Singapore, 843–860. https://doi.org/10.1007/978-981-10-5251-4_103

[7] Alberto Cairo. 2019. *How Charts Lie: Getting Smarter about Visual Information*. W. W. Norton & Company, United States.

[8] Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. 2011. Information Credibility on Twitter. In *Proceedings of the 20th International Conference on World Wide Web*. ACM, Hyderabad India, 675–684. https://doi.org/10.1145/1963405.1963500

[9] Quan Ze Chen and Amy X. Zhang. 2023. Judgment Sieve: Reducing Uncertainty in Group Judgments through Interventions Targeting Ambiguity versus Disagreement. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–26. https://doi.org/10.1145/3610074

[10] Aaron Clauset, Cosma Rohilla Shalizi, and M. E. J. Newman. 2009. Power-Law Distributions in Empirical Data. *SIAM Rev.* 51, 4 (2009), 661–703. https://doi.org/10.1137/070710111

[11] John Cook, Naomi Oreskes, Peter T Doran, William R L Anderegg, Bart Verheggen, Ed W Maibach, J Stuart Carlton, Stephan Lewandowsky, Andrew G Skuce, Sarah A Green, Dana Nuccitelli, Peter Jacobs, Mark Richardson, Bärbel Winkler, Rob Painting, and Ken Rice. 2016. Consensus on Consensus: A Synthesis of Consensus Estimates on Human-Caused Global Warming. *Environmental Research Letters* 11, 4 (2016), 048002. https://doi.org/10.1088/1748-9326/11/4/048002

[12] Matthew N. Hannah. 2021. A Conspiracy of Data: QAnon, Social Media, and Information Visualization. *Social Media + Society* 7, 3 (2021), 1–15. https://doi.org/10.1177/20563051211036064

[13] Sun-ha Hong. 2023. Fact Signalling and Fact Nostalgia in the Data-Driven Society. *Big Data & Society* 10, 1 (2023), 1–12. https://doi.org/10.1177/20539517231164118

[14] Darrell Huff. 1993. *How to Lie with Statistics*. W. W. Norton & Company, United States.

[15] Jessica Hullman, Eytan Adar, and Priti Shah. 2011. The Impact of Social Information on Visual Judgments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Vancouver BC Canada, 1461–1470. https://doi.org/10.1145/1978942.1979157

[16] Farnaz Jahanbakhsh, Amy X. Zhang, Adam J. Berinsky, Gordon Pennycook, David G. Rand, and David R. Karger. 2021. Exploring Lightweight Interventions at Posting Time to Reduce the Sharing of Misinformation on Social Media. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW (2021), 1–42. https://doi.org/10.1145/3449092

[17] Farnaz Jahanbakhsh, Amy X. Zhang, and David R. Karger. 2022. Leveraging Structured Trusted-Peer Assessments to Combat Misinformation. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW (2022), 1–40. https://doi.org/10.1145/3555637

[18] Nigel King. 2004. Using Templates in the Thematic Analysis of Text. In *Essential Guide to Qualitative Methods in Organizational Research*, Catherine Cassell and Gillian Symon (Eds.). SAGE Publications Ltd, London, UK, 256–270. https://doi.org/10.4135/9781446280119

[19] Ha-Kyung Kong, Zhicheng Liu, and Karrie Karahalios. 2018. Frames and Slants in Titles of Visualizations on Controversial Topics. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–12. https://doi.org/10.1145/3173574.3174012

[20] Thomas S. Kuhn and Ian Hacking. 2012. *The Structure of Scientific Revolutions* (4th ed.). The University of Chicago Press, Chicago ; London.

[21] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. 2010. What Is Twitter, a Social Network or a News Media?. In *Proceedings of the 19th International Conference on World Wide Web (WWW '10)*. Association for Computing Machinery, New York, NY, USA, 591–600. https://doi.org/10.1145/1772690.1772751

[22] Bruno Latour. 2015. *Science in Action: How to Follow Scientists and Engineers through Society* (nachdr. ed.). Harvard Univ. Press, Cambridge, Mass.

[23] Crystal Lee, Tanya Yang, Gabrielle D Inchoco, Graham M. Jones, and Arvind Satyanarayan. 2021. Viral Visualizations: How Coronavirus Skeptics Use Orthodox Data Practices to Promote Unorthodox Science Online. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, 607:1–607:18.

[24] Yiyi Li and Ying Xie. 2020. Is a Picture Worth a Thousand Words? An Empirical Study of Image Content and Social Media Engagement. *Journal of Marketing Research* 57, 1 (2020), 1–19. https://doi.org/10.1177/0022243719881113

[25] Cochrane Library. 2023. About the Cochrane Database of Systematic Reviews | Cochrane Library. https://www.cochranelibrary.com/cdsr/about-cdsr.

[26] Moshe Lichman and Padhraic Smyth. 2018. Prediction of Sparse User-Item Consumption Rates with Zero-Inflated Poisson Regression. In *Proceedings of the 2018 World Wide Web Conference (WWW '18)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 719–728. https://doi.org/10.1145/3178876.3186153

[27] Haihan Lin, Derya Akbaba, Miriah Meyer, and Alexander Lex. 2023. Data Hunches: Incorporating Personal Knowledge into Visualizations. *IEEE Transactions on Visualization and Computer Graphics* 29, 1 (2023), 504–514. https://doi.org/10.1109/TVCG.2022.3209451

[28] Maxim Lisnic, Cole Polychronis, Alexander Lex, and Marina Kogan. 2023. Misleading Beyond Visual Tricks: How People Actually Lie with Charts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg Germany, 1–21. https://doi.org/10.1145/3544548.3580910

[29] Leo Yu-Ho Lo, Ayush Gupta, Kento Shigyo, Aoyu Wu, Enrico Bertini, and Huamin Qu. 2022. Misinformed by Visualization: What Do We Learn From Misinformative Visualizations? *Computer Graphics Forum* 41, 3 (2022), 515–525. https://doi.org/10.1111/cgf.14559

[30] Hana Matatov, Mor Naaman, and Ofra Amir. 2022. Stop the [Image] Steal: The Role and Dynamics of Visual Content in the 2020 U.S. Election Misinformation Campaign. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 541:1–541:24. https://doi.org/10.1145/3555599

[31] Edouard Mathieu, Hannah Ritchie, Lucas Rodés-Guirao, Cameron Appel, Charlie Giattino, Joe Hasell, Bobbie Macdonald, Saloni Dattani, Diana Beltekian, Esteban Ortiz-Ospina, and Max Roser. 2020. Coronavirus Pandemic (COVID-19). https://ourworldindata.org/coronavirus.

[32] Andrew McNutt, Gordon Kindlmann, and Michael Correll. 2020. Surfacing Visualization Mirages. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–16. https://doi.org/10.1145/3313831.3376420

[33] Mohsen Mosleh, Cameron Martel, Dean Eckles, and David Rand. 2021. Perverse Downstream Consequences of Debunking: Being Corrected by Another User for Posting False Political News Increases Subsequent Sharing of Low Quality, Partisan, and Toxic Content in a Twitter Field Experiment. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–13. https://doi.org/10.1145/3411764.3445642

[34] Naomi Oreskes. 2004. The Scientific Consensus on Climate Change. *Science* 306, 5702 (2004), 1686–1686. https://doi.org/10.1126/science.1103618

[35] Naomi Oreskes and Erik M. Conway. 2010. *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming* (1st ed.). Bloomsbury Press, New York.

[36] Anshul Vikram Pandey, Katharina Rall, Margaret L. Satterthwaite, Oded Nov, and Enrico Bertini. 2015. How Deceptive Are Deceptive Visualizations? An Empirical Analysis of Common Distortion Techniques. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. Association for Computing Machinery, Seoul, Republic of Korea, 1469–1478. https://doi.org/10.1145/2702123.2702608

[37] Orestis Papakyriakopoulos and Ellen Goodman. 2022. The Impact of Twitter Labels on Misinformation Spread and User Engagement: Lessons from Trump's Election Tweets. In *Proceedings of the ACM Web Conference 2022*. ACM, Virtual Event, Lyon France, 2541–2551. https://doi.org/10.1145/3485447.3512126

[38] Evan M. Peck, Sofia E. Ayuso, and Omar El-Etr. 2019. Data Is Personal: Attitudes and Perceptions of Data Visualization in Rural Pennsylvania. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300474

[39] Gordon Pennycook, Adam Bear, Evan T. Collins, and David G. Rand. 2020. The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Headlines Increases Perceived Accuracy of Headlines Without Warnings. *Management Science* 66, 11 (2020), 4944–4957. https://doi.org/10.1287/mnsc.2019.3478

[40] Uri Shwed and Peter S. Bearman. 2010. The Temporal Structure of Scientific Consensus Formation. *American Sociological Review* 75, 6 (2010), 817–840. https://doi.org/10.1177/0003122410388488

[41] Ian Skurnik, Carolyn Yoon, Denise C. Park, and Norbert Schwarz. 2005. How Warnings about False Claims Become Recommendations. *Journal of Consumer Research* 31, 4 (2005), 713–724. https://doi.org/10.1086/426605

[42] Sahana Sule, Marisa C. DaCosta, Erin DeCou, Charlotte Gilson, Kate Wallace, and Sarah L. Goff. 2023. Communication of COVID-19 Misinformation on Social Media by Physicians in the US. *JAMA Network Open* 6, 8 (2023), e2328928. https://doi.org/10.1001/jamanetworkopen.2023.28928

[43] Eric Sun, Itamar Rosenn, Cameron Marlow, and Thomas Lento. 2009. Gesundheit! Modeling Contagion through Facebook News Feed. *Proceedings of the International AAAI Conference on Web and Social Media* 3, 1 (2009), 146–153. https://doi.org/10.1609/icwsm.v3i1.13947

[44] Joseph E. Uscinski and Ryden W. Butler. 2013. The Epistemology of Fact Checking. *Critical Review* 25, 2 (2013), 162–180. https://doi.org/10.1080/08913811.2013.843872

[45] Meredith Wadman. 2021. Antivaccine Activists Use a Government Database on Side Effects to Scare the Public. https://www.sciencemag.org/news/2021/05/antivaccine-activists-use-government-database-side-effects-scare-public.

[46] Nathan Walter, Jonathan Cohen, R. Lance Holbert, and Yasmin Morag. 2020. Fact-Checking: A Meta-Analysis of What Works and for Whom. *Political Communication* 37, 3 (2020), 350–375. https://doi.org/10.1080/10584609.2019.1668894

[47] Nathan Walter and Nikita A. Salovich. 2021. Unchecked vs. Uncheckable: How Opinion-Based Claims Can Impede Corrections of Misinformation. *Mass Communication and Society* 24, 4 (2021), 500–526. https://doi.org/10.1080/15205436.2020.1864406

[48] Teresa Weikmann and Sophie Lecheler. 2022. Visual Disinformation in a Digital Age: A Literature Synthesis and Research Agenda. *New Media & Society* 25, 12 (2022), 146144482211416. https://doi.org/10.1177/14614448221141648

[49] Worldometers. 2023. COVID — Coronavirus Statistics — Worldometer. https://www.worldometers.info/coronavirus/.

[50] Cindy Xiong, Lisanne Van Weelden, and Steven Franconeri. 2020. The Curse of Knowledge in Visual Data Communication. *IEEE Transactions on Visualization and Computer Graphics* 26, 10 (2020), 3051–3062. https://doi.org/10.1109/TVCG.2019.2917689

[51] Elizabeth Zak. 2023. The Colors of a #climatescam: An Exploration of Anti-Climate Change Graphs on Twitter. *Journal of Interdisciplinary Sciences* 7, 1 (2023), 13–28.

[52] Savvas Zannettou. 2021. "I Won the Election!": An Empirical Analysis of Soft Moderation Interventions on Twitter. *Proceedings of the International AAAI Conference on Web and Social Media* 15 (2021), 865–876. https://doi.org/10.1609/icwsm.v15i1.18110

[53] Amy X. Zhang, Martin Robbins, Ed Bice, Sandro Hawke, David Karger, An Xiao Mina, Aditya Ranganathan, Sarah Emlen Metz, Scott Appling, Connie Moon Sehat, Norman Gilmore, Nick B. Adams, Emmanuel Vincent, and Jennifer Lee. 2018. A Structured Response to Misinformation: Defining and Annotating Credibility Indicators in News Articles. In *Companion of the The Web Conference 2018 on The Web Conference 2018 - WWW '18*. ACM Press, Lyon, France, 603–612. https://doi.org/10.1145/3184558.3188731

[54] C. Ziemkiewicz, A. Ottley, R. J. Crouser, K. Chauncey, S. L. Su, and R. Chang. 2012. Understanding Visualization by Understanding Individual Users. *IEEE Computer Graphics and Applications* 32, 6 (2012), 88–94. https://doi.org/10.1109/MCG.2012.120