

“Perfect moderation does not exist,” but here are some lessons from Twitter’s labels of Trump tweets | Nieman Journalism Lab

By Shraddha Chakradhar

March 13, 2022

From simple labeling to links to third-party sites for accurate information to outright blocking, social media platforms continue to test different ways to keep users informed about content containing mis- and disinformation. A lot of these efforts — often, by the companies’ own admission — have proven less than fruitful.

Now, a new study that analyzed former president Donald Trump’s tweets finds that there are no clear-cut answers when it comes to determining how people engage with content that’s clearly labeled as containing misleading information. The study drew from more than 1,200 tweets from Donald Trump’s Twitter account between October 30, 2020 and January 8, 2021. The authors also collected these tweets’ engagement metrics as well as the nearly 2.4 million replies to these tweets.

At the outset, the study found that labeled tweets were much more likely to garner user engagement: Compared to unlabeled tweets, labeled tweets were liked approximately 36% more, retweeted 70% more, quote tweeted 88% more, and the median number of replies labeled tweets generated was 84% higher than that of unlabeled tweets.

[...]

Source: [“Perfect moderation does not exist,” but here are some lessons from Twitter’s labels of Trump tweets | Nieman Journalism Lab](#)