

# A Comparison of Account-Focused and Content-Focused Warnings on User Trust of Twitter Content

Ronald Thompson, Santana Koring'ura, Marshini Chetty and Daniel Votipka

## Research Questions

Are account-focused warnings more effective at stopping the spread of disinformation?

What effect does political leaning have on comfort with warnings on Twitter?

## Warnings Tested

Account Connections – account is connected to a disinformation campaign

Account Behavior – account is misrepresenting themselves [see example]

Qualitative Evaluation – expert panel has connected account to disinformation campaign

Twitter – current warning Twitter uses for content determined to be disinformation

## Recommendations

1. Account-focused warnings are an effective tool for slowing possible disinformation
2. Warnings are not a silver bullet & considered in context of a broader systemic change

## Survey Variables

Veracity	Partisanship	Warning	Quantitative
True	Republican	No warning	
False	Democrat	Connections	True
	Public Health	Behavior	False
		Qualitative	
		Twitter	

## Example Tweet & Warning

**Quoting Account Profile Picture**

**Quoting Account Name**

Medicare 4 All @medicare4all

This account might be misleading about an election or other civic process. Our systems have classified, with 73% certainty, the account owner may be misrepresenting who they are or is acting like an automated bot. This behavior is common in disinformation campaigns. [Learn more](#) **Warning**

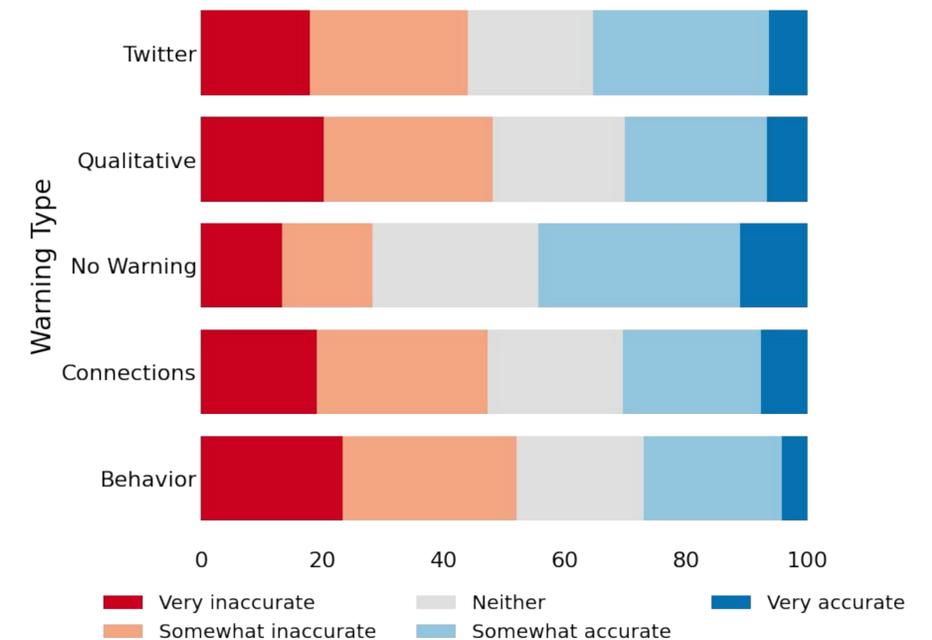
**The Chronicle** @TheChronicle · 6h **Original Account**

Republican School Board Member pushing to teach Civil War only from the perspective of slave owners <https://bit.ly/2UplIjd>

4:39 PM · 7/5/21 · [Twitter Web App](#) **Original Tweet**

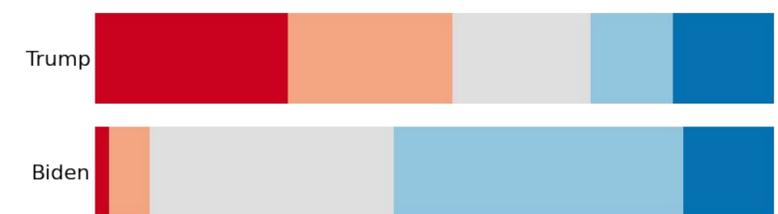
## Warning Effectiveness

How accurate is the claim in the headline?



## Warning Comfort

How comfortable are you with ACCOUNT warnings?



How comfortable are you with CONTENT warnings?

