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

Tufts
UNIVERSITY

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	Connections	True
False	Democrat	Behavior	False
	Public Health	Qualitative	
		Twitter	
		No warning	

Example Tweet & Warning

Quoting Account Profile Picture

Medicare 4 All
@medicare4all **Quoting Account Name**

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/2UpllJd

4:39 PM · 7/5/21 · Twitter Web App **Original Tweet**

Warning Effectiveness

How accurate is the claim in the headline?

Legend: Very inaccurate · Somewhat inaccurate · Neither · Somewhat accurate · Very accurate
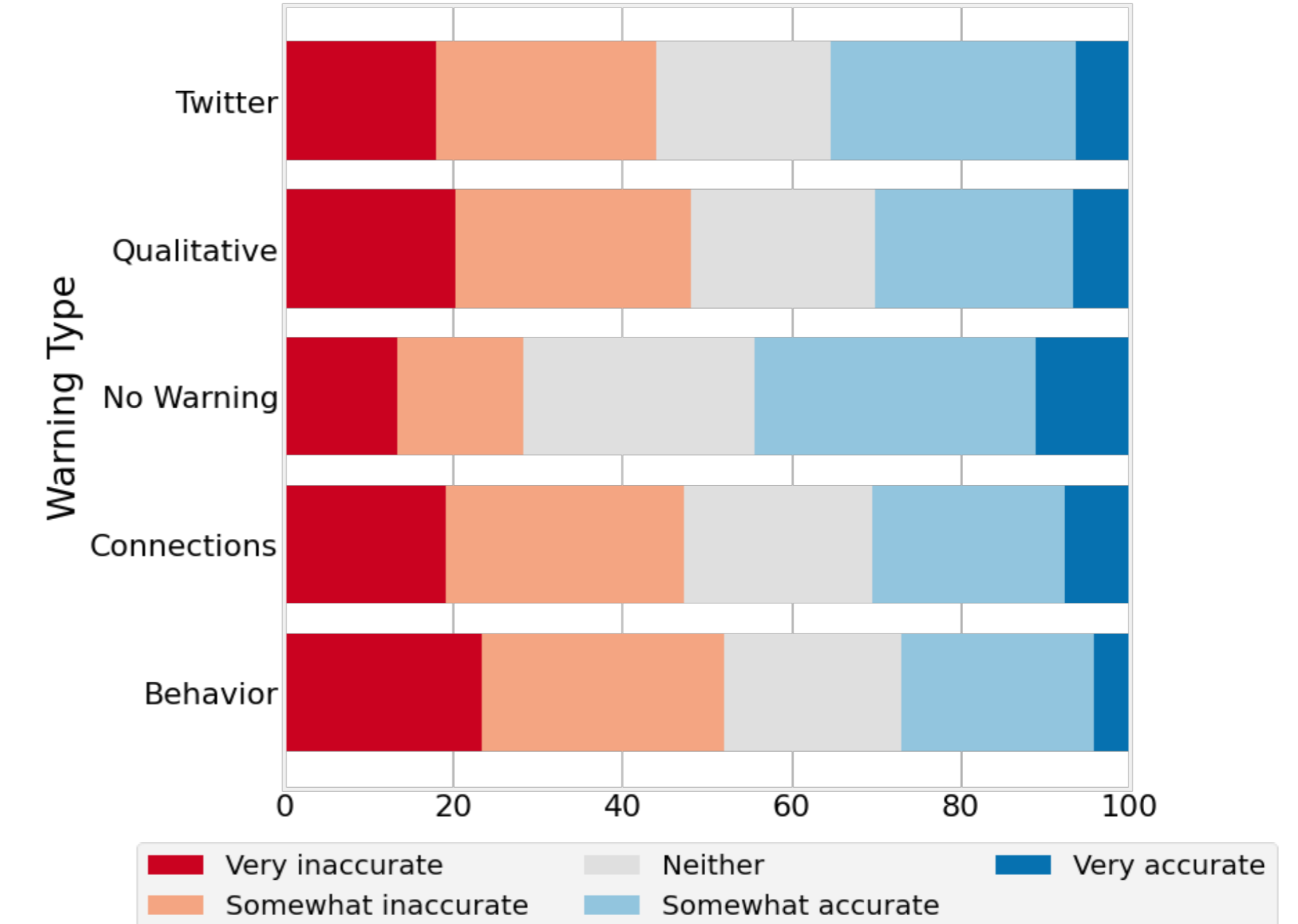
Warning Comfort

How comfortable are you with ACCOUNT warnings?
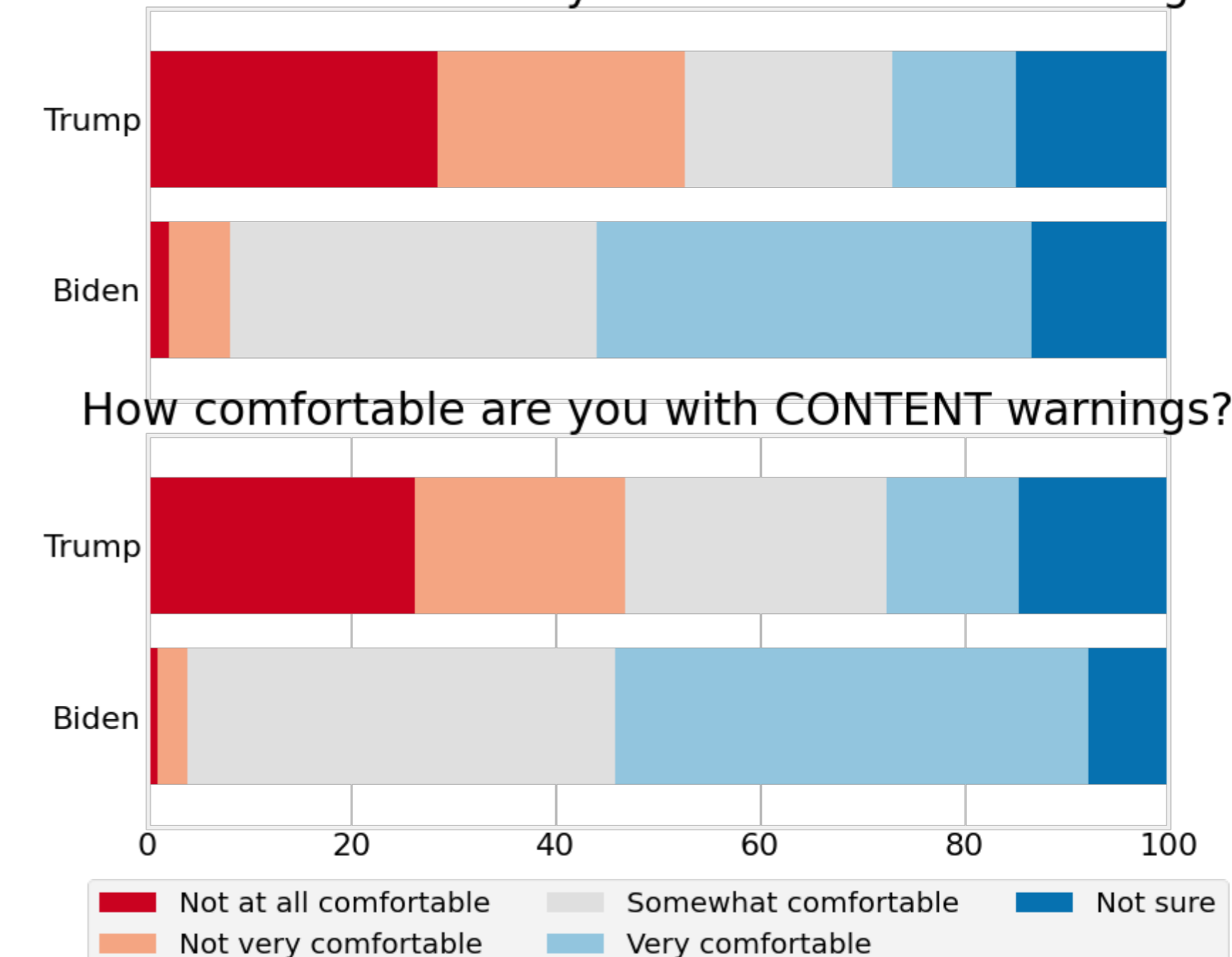
How comfortable are you with CONTENT warnings?

Legend: Not at all comfortable · Not very comfortable · Somewhat comfortable · Very comfortable · Not sure