

“I can say I’m John Travolta...but I’m not John Travolta”: Investigating the Impact of Changes to Social Media Verification Policies on User Perceptions of Verified Accounts



Carson Powers, Nickolas Gravel, Christopher Pellegrini
Micah Sherr, Michelle Mazurek, Daniel Votipka

Tufts Security & Privacy Lab





Background – Account Verification

- Almost all social media and microblogging sites support *account verification*
- Denoted with an indicator after the account name

Instagram: tuftsuniversity 

LinkedIn: Alex Olsen 

Twitter: Twitter Verified 

Snapchat: Ariana Grande 




Background – Account Verification

- Almost all social media and microblogging sites support *account verification*
- Denoted with an indicator after the account name
- Historically reserved for *notable individuals* (e.g., celebrities, journalists, government institutions, etc.)

Instagram: tuftsuniversity 

Twitter: Twitter Verified 

LinkedIn: Alex Olsen 

Snapchat: Ariana Grande 




Background – Account Verification

- Almost all social media and microblogging sites support *account verification*
- Denoted with an indicator after the account name
- Historically reserved for *notable individuals* (e.g., celebrities, journalists, government institutions, etc.)
- Individual was authenticated with government ID

Instagram: tuftsuniversity 

Twitter: Twitter Verified 

LinkedIn: Alex Olsen 

Snapchat: Ariana Grande 

Background – Perception of Verification



- Authenticates the account's identity and does not endorse validity
- “Verified” does not necessarily mean “true”...but do people see it that way?
 - Users understand the verified indicator confers authenticity, not credibility [Vaidya et al. 2019]



Elon Musk ✓ 

@elonmusk



Twitter's current lords & peasants system for who has or doesn't have a blue checkmark is bullshit.

Power to the people! Blue for \$8/month.

13:36 · 01.11.22



Elon Musk ✓ 

@elonmusk



Twitter's current lords & peasants system for who has or doesn't have a blue checkmark is bullshit.

Power to the people! Blue for \$8/month.

13:36 · 01.11.22



Mark Zuckerberg ✓



53 m · 🌐

Good morning and new product announcement: this week we're starting to roll out Meta Verified -- a subscription service that lets you verify your account with a government ID, get a blue badge, get extra impersonation protection against accounts claiming to be you, and get direct access to customer support. This new feature is about increasing authenticity and security across our services. Meta Verified starts at \$11.99 / month on web or \$14.99 / month on iOS. We'll be rolling out in Australia and New Zealand this week and more countries soon.

Our new verification features

Published on Apr 12, 2023 | Categories: [Product News](#)



LinkedIn Corporate Communications



Today we announced new, free ways people who use LinkedIn can verify their identity, work email addresses and workplaces on LinkedIn. You can read more about the announcement [here](#) by Oscar Rodriguez, our head of trust and privacy.

As these new features and tools roll out, millions of people who use LinkedIn will now have the ability to

- Verify their identity with CLEAR
- Verify their work email address
- Verify their workplace through Microsoft Entra

We're excited to be partnering with industry leaders Microsoft Entra and CLEAR to help ensure that LinkedIn is your source for trusted, verified information.

Background



We cannot assume that previously observed perceptions of verified accounts are still valid.

Research Questions



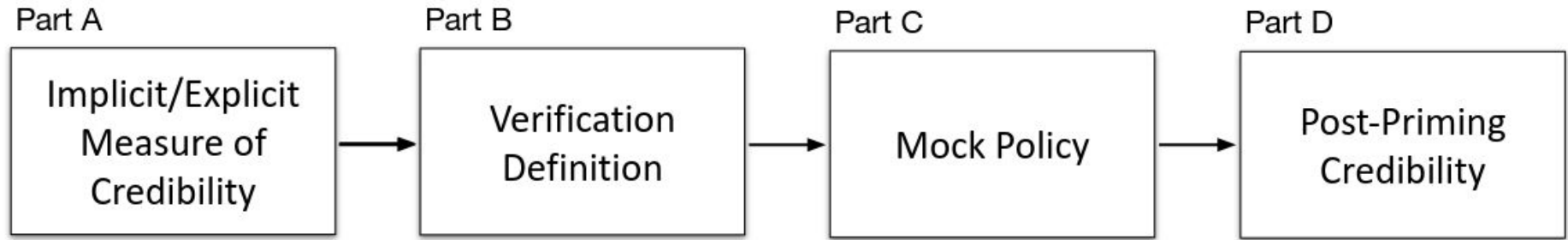
- 1) What are the verification policies used by popular social media platforms and how have they changed over time?
- 2) What do users think account verification entails? How does it impact perceptions of posted content credibility?
- 3) How would potential changes to verification policies impact user perceptions of posts from verified accounts and user perceptions of the policies?

Research Questions

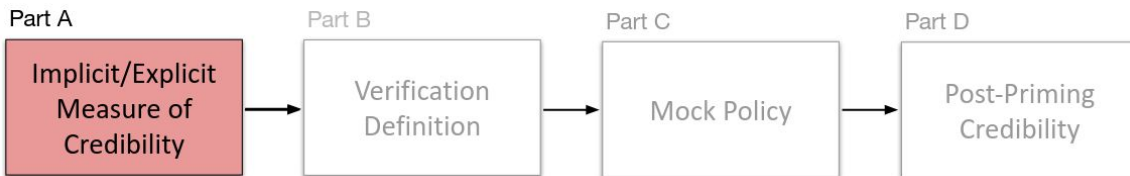


- 1) What are the verification policies used by popular social media platforms and how have they changed over time?
- 2) What do users think account verification entails? How does it impact perceptions of posted content credibility?
- 3) How would potential changes to verification policies impact user perceptions of posts from verified accounts and user perceptions of the policies?

Survey Design



Example post



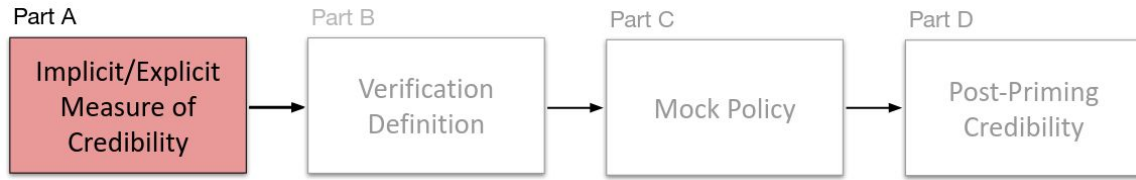
Post A



Post B

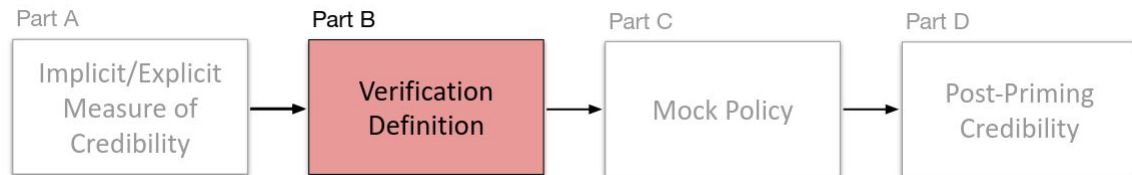
1. Post A and Post B contain conflicting information. Which of the posts do you believe is correct? (5-pt Likert scale “Definitely A” to “Definitely B”)

Conditions



- Posts were formatted to look like either *Facebook* or *Twitter* (randomly assigned)
- Participants saw either a post about the connection between coffee and disease, or a post about a suspected bomb threat (*Coffee* or *Police*)

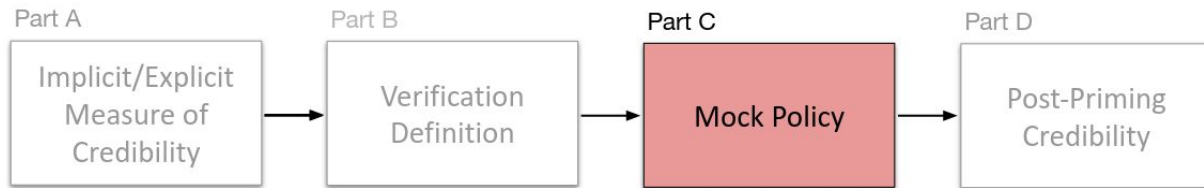
Verification Definition



One of the tweets you were previously shown was by an account with a verification checkmark (✓) indicating that the account has been verified.

Based on your understanding of [Twitter's/Facebook's] account verification, what requirements must an account satisfy to become verified and obtain a verified checkmark?

Policy Acceptability



Suppose Twitter adopted a verification policy in which the account had to meet all of the following criteria:

- **Any user** on the platform is allowed to apply for verification
- Accounts must submit **government-issued identification** that matches the name of the account being verified
- Accounts **pay a monthly subscription fee** to maintain their verification checkmark

1. To what level do you believe these verification requirements are acceptable for verifying account owner identity?
 - (a) Unacceptable
 - (b) Slightly Unacceptable
 - (c) Neutral
 - (d) Slightly Acceptable
 - (e) Acceptable
2. If you could suggest one thing to add, remove, or change in this policy to improve its ability in verifying the account owner is who they say they are, what would it be? Please explain why



Mock Policy Dimensions

Eligibility

- **Notable:** Verification is for persons of interest only
- **Open:** Anyone can be verified

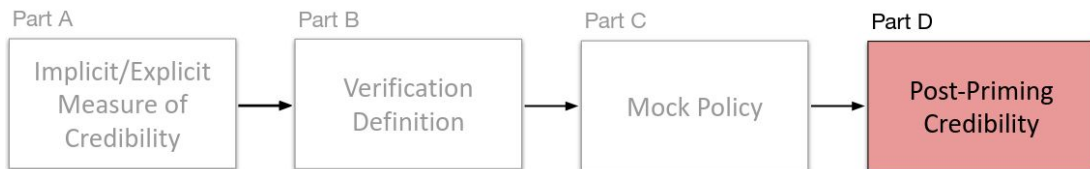
Verification Method

- **Government ID:** User must prove identity with government-issued ID
- **Phone/Email:** User must verify their phone number or email address

Payment

- **Paid:** User must pay monthly fee to maintain verified status
- **Free:** No payment required

Credibility After Priming



Post A



Post B

- Participants were again asked which post was correct *assuming the presented policy was used for verification*
- Measured whether the presented policy affected perceived credibility

Research Questions



- 1) What are the verification policies used by popular social media platforms and how have they changed over time?
- 2) What do users think account verification entails? How does it impact perceptions of posted content credibility?
- 3) How would potential changes to verification policies impact user perceptions of posts from verified accounts and user perceptions of the policies?

Differences in Verification Policies



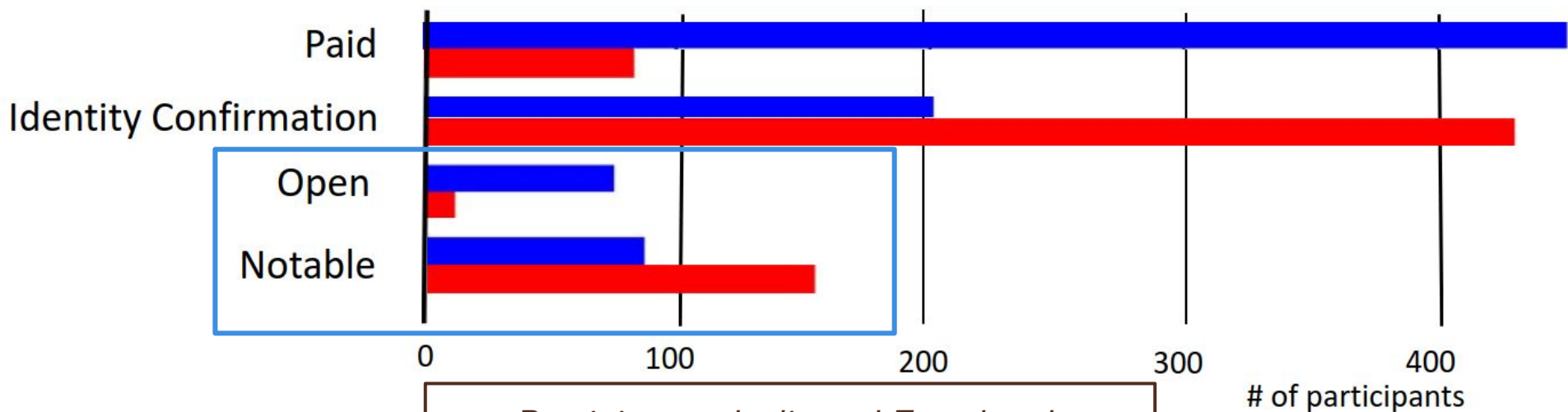
	Twitter	Facebook
Eligibility	Open	Open
Verification Method	Phone Number	Gov ID
Payment	Paid	Paid

Results – Definition of Verification



- Twitter
- Facebook

Definition of Verification



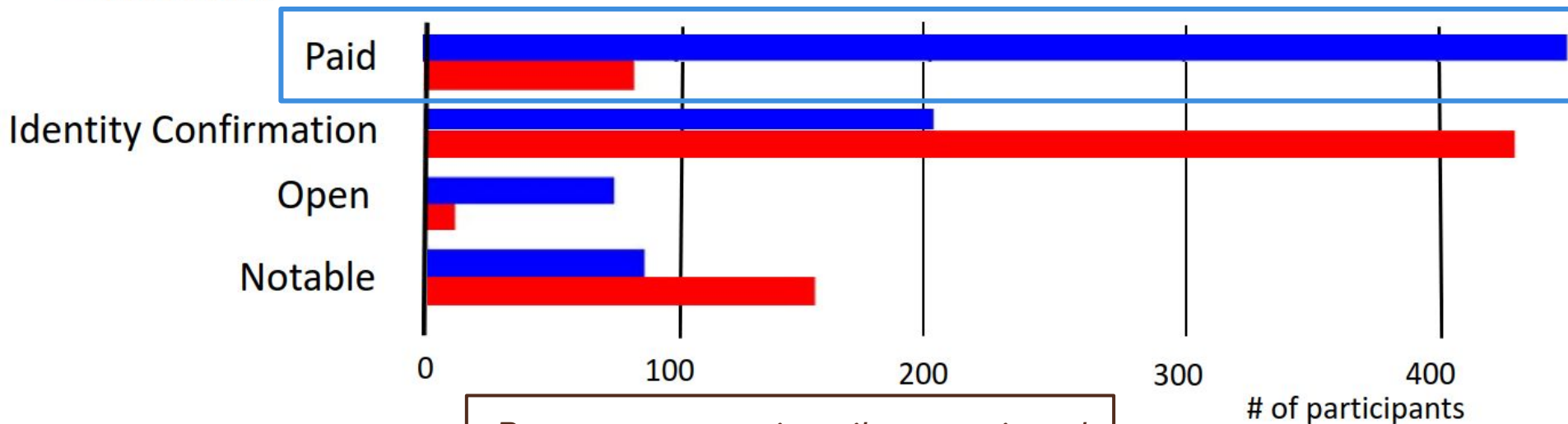
Participants believed Facebook verification is for notable only

Results – Definition of Verification



- Twitter
- Facebook

Definition of Verification



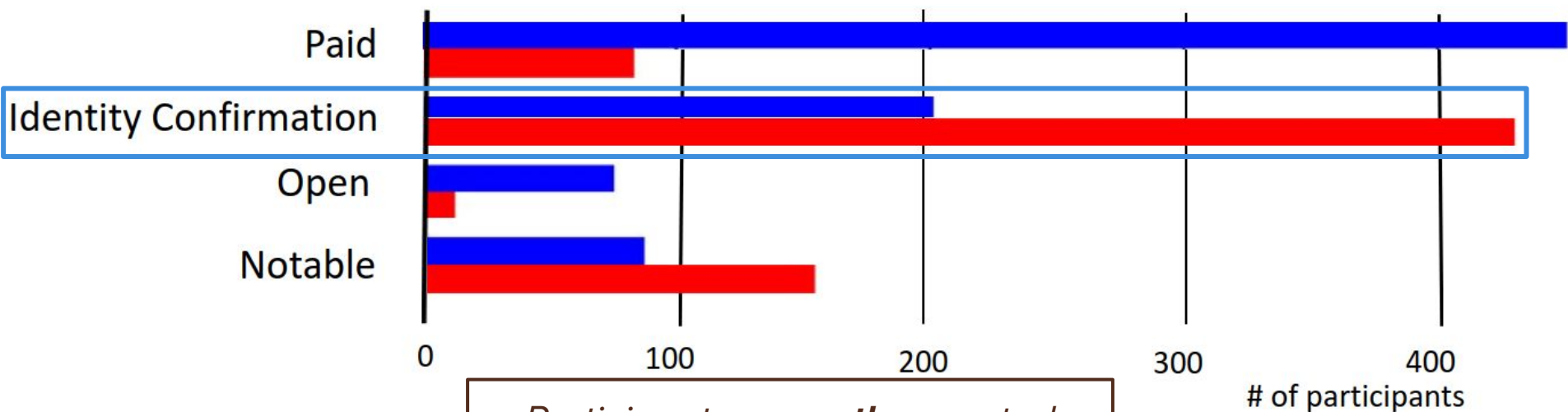
Payment was primarily associated with Twitter

Results – Definition of Verification



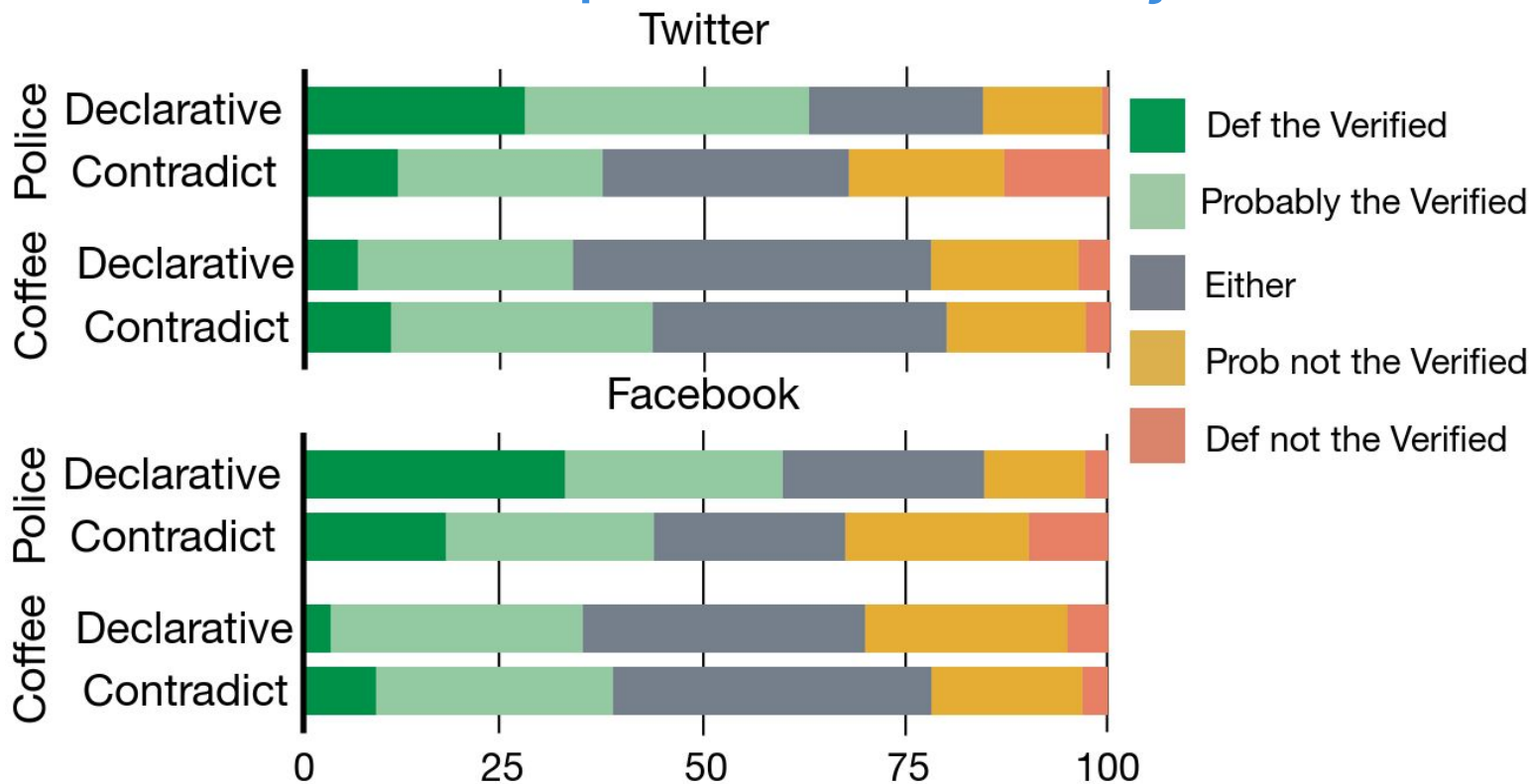
- Twitter
- Facebook

Definition of Verification



*Participants **correctly** reported
Facebook confirms identity*

Platform did not affect perceived credibility

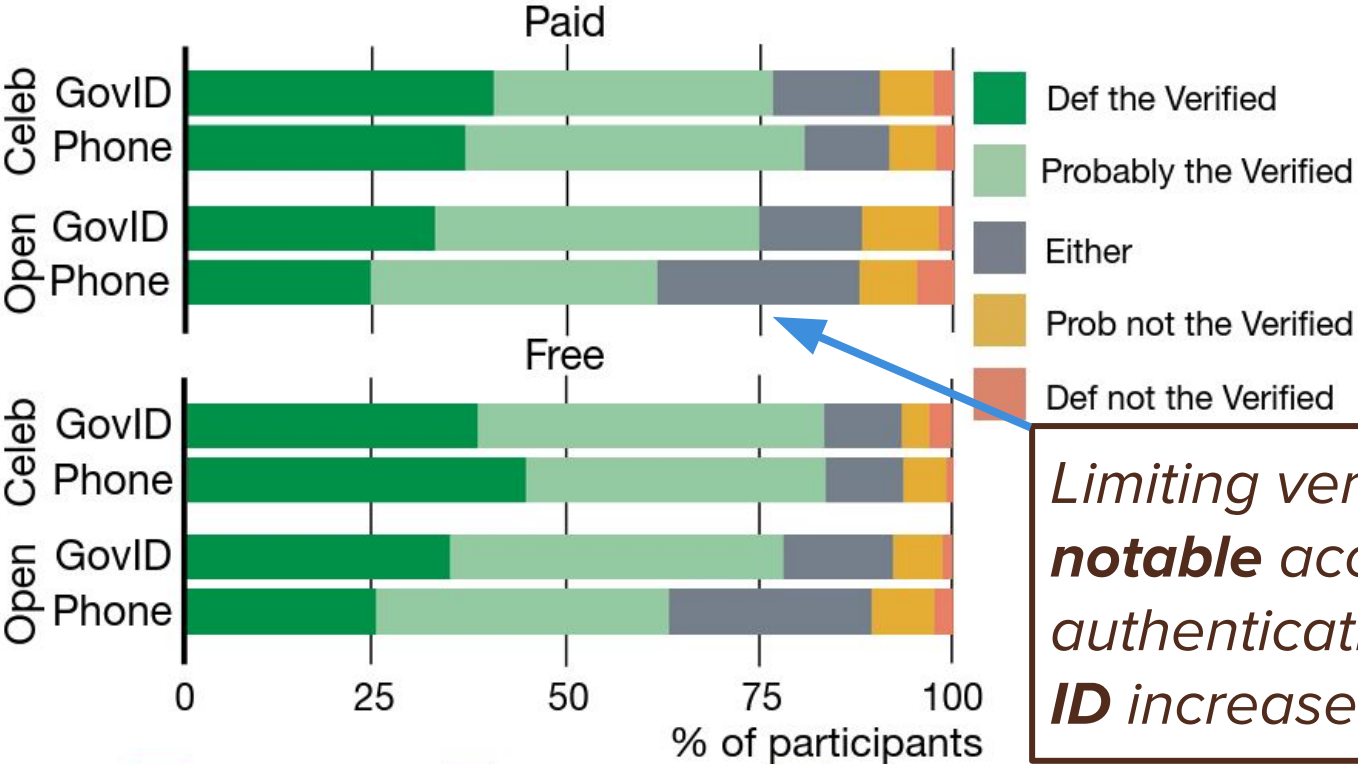


Research Questions



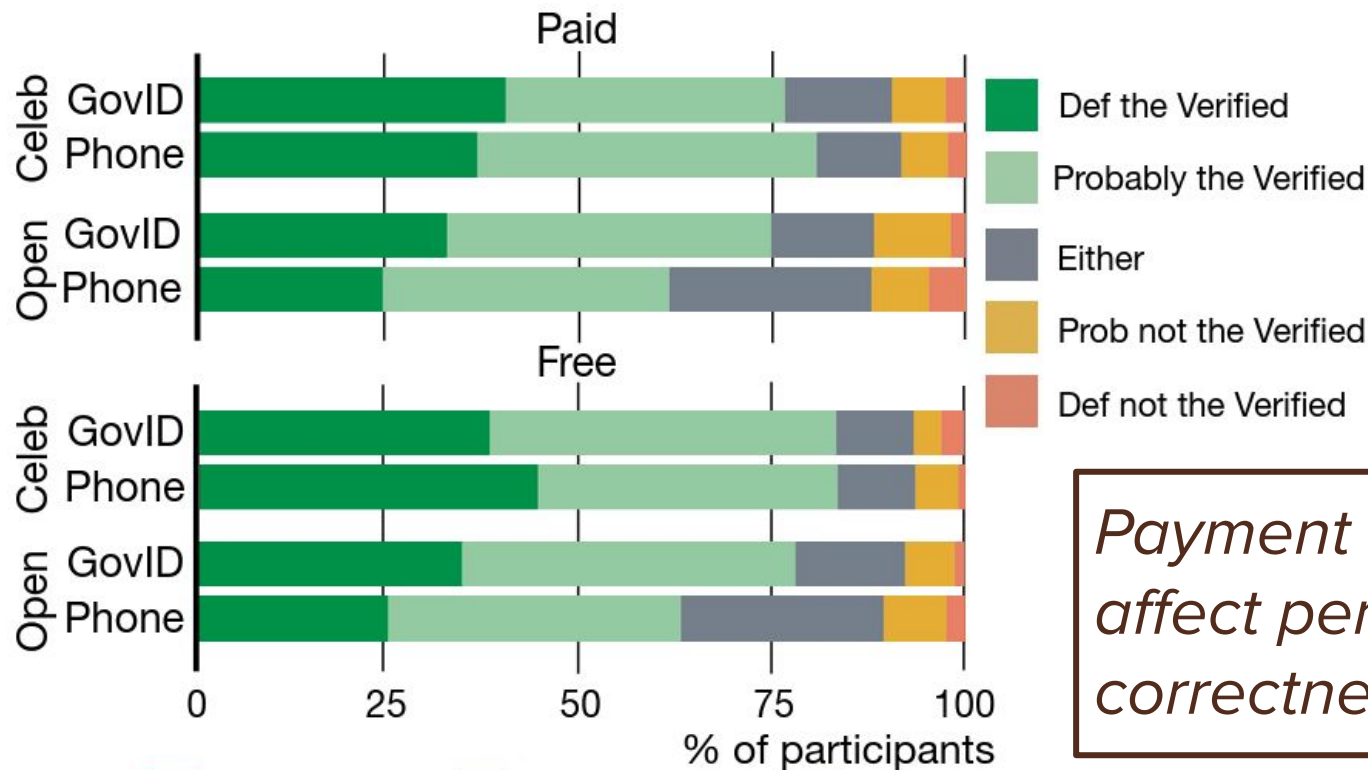
- 1) What are the verification policies used by popular social media platforms and how have they changed over time?
- 2) What do users think account verification entails? How does it impact perceptions of posted content credibility?
- 3) How would potential changes to verification policies impact user perceptions of posts from verified accounts and user perceptions of the policies?

Mock Policy's Effect on Content Credibility Perception



*Limiting verification to **notable** accounts and authenticating with **Gov ID** increased credibility*

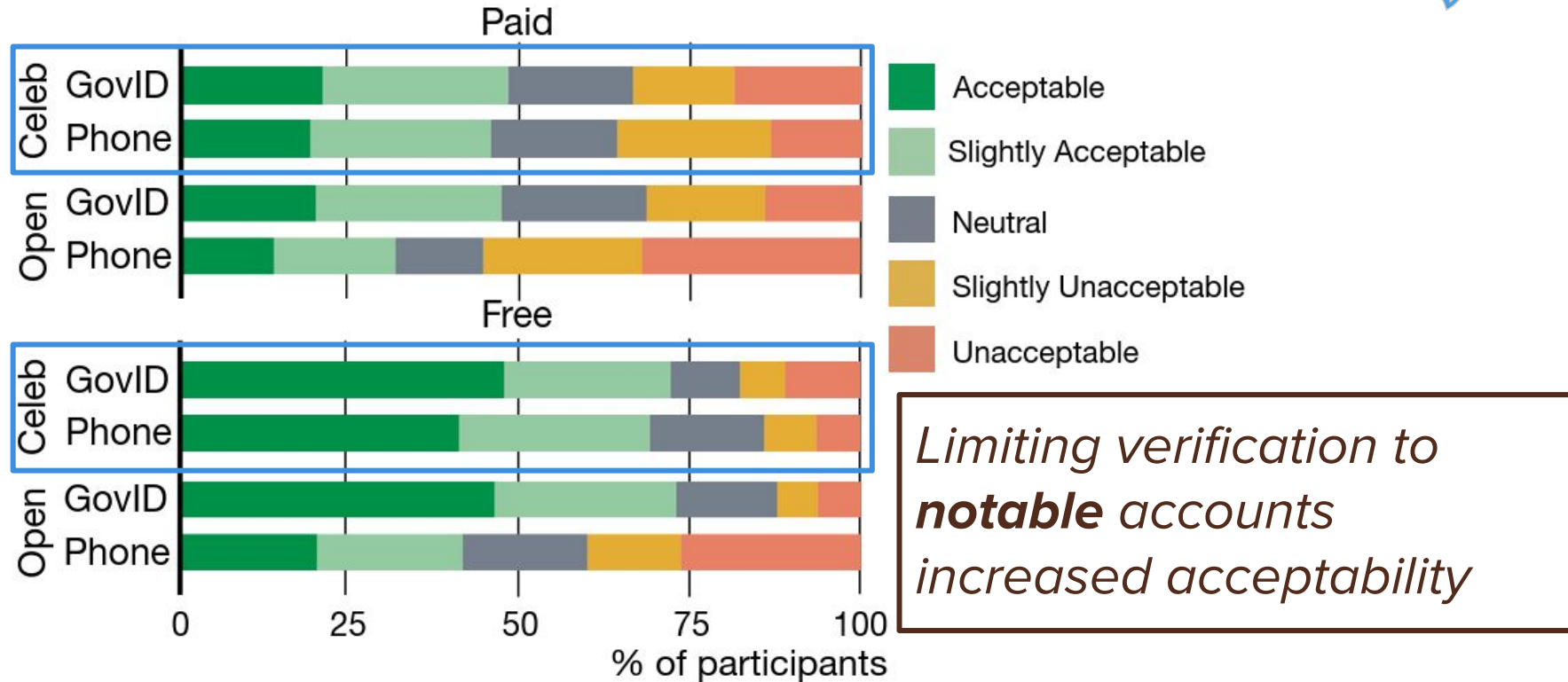
Mock Policy's Effect on Content Credibility Perception



Payment does not affect perceived correctness



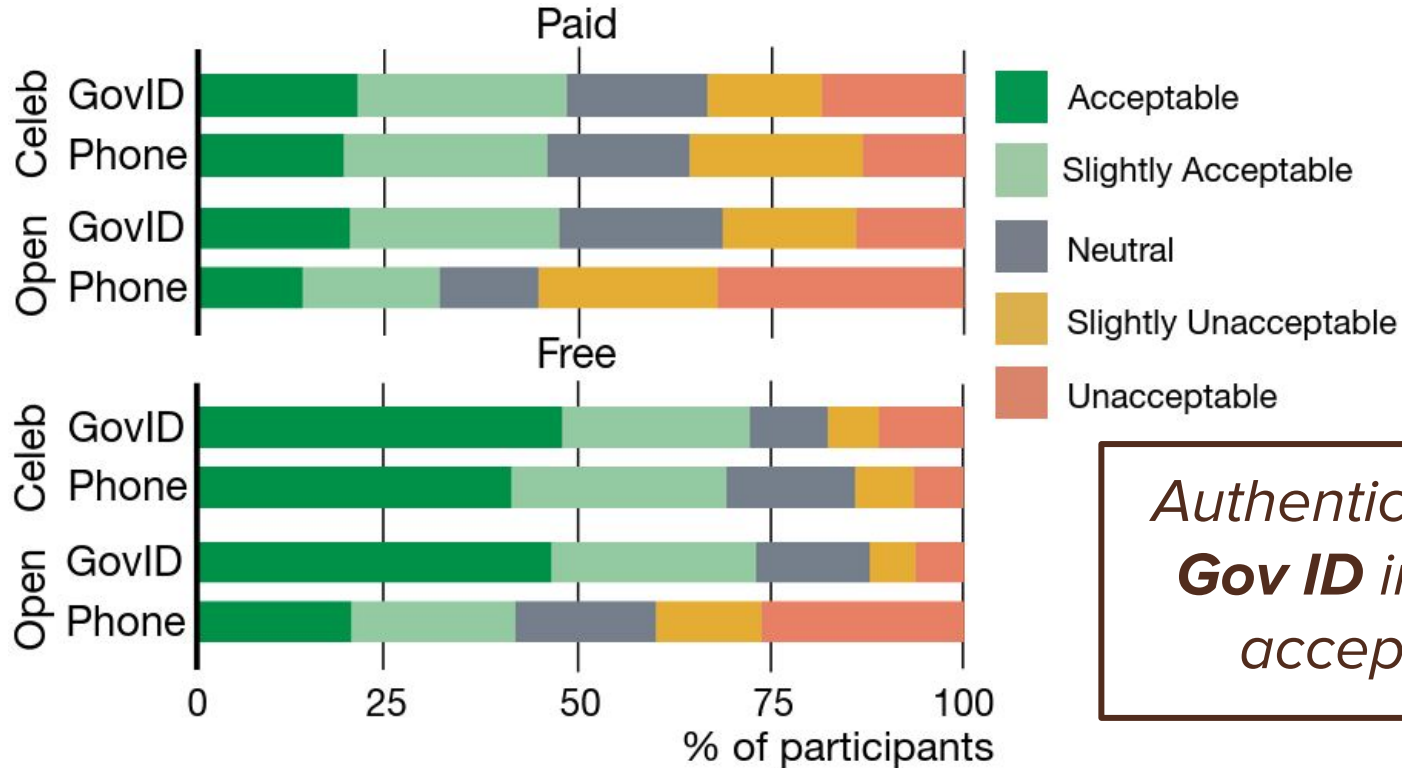
Perception of Mock Verification Policies



*Limiting verification to **notable** accounts increased acceptability*



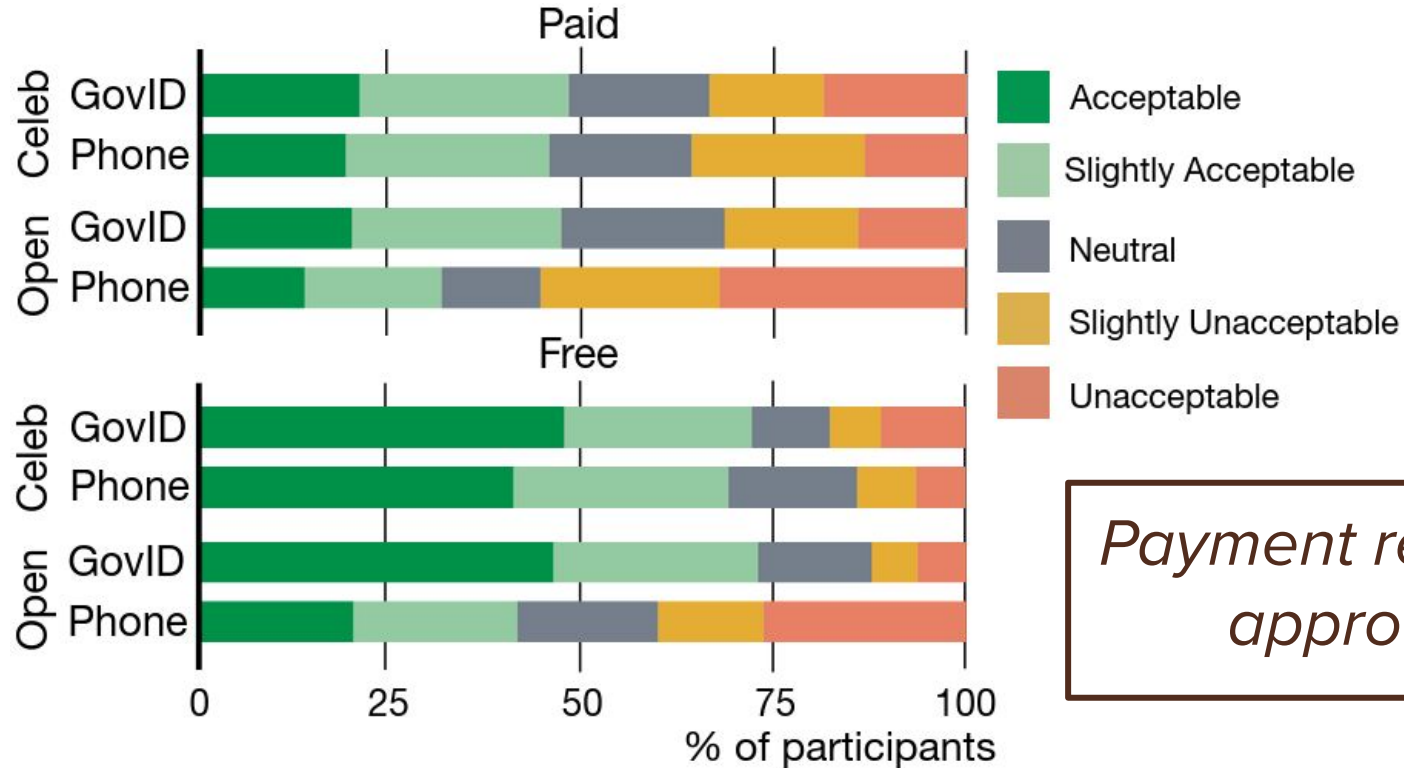
Perception of Mock Verification Policies



*Authenticating with
Gov ID increased
acceptability*



Perception of Mock Verification Policies



Payment reduced approval



Conclusion

- Participants were more aware of Twitter’s policy changes than Facebook’s
- No difference in the verified indicator’s effect between platforms
- Participants preferred government ID-authenticated, free policies

Recommendations:

- Platforms should require government ID checks to improve trust in verification
- “Tiered verification” may resolve the split between wanting open or notable-only eligibility

Conclusion

Questions:

carson.powers@tufts.edu

tsp.cs.tufts.edu



- Participants were more aware of Twitter’s policy changes than Facebook’s
- No difference in the verified indicator’s effect between platforms
- Participants preferred government ID-authenticated, free policies

Recommendations:

- Platforms should include government ID checks to improve trust in verification
- “Tiered verification” may resolve the split between wanting open or notable-only eligibility