facebook
Product Policy Forum
May 7, 2019

TOPICS: Violence and Graphic Content Interstitials, Voter and Census Interference
<table>
<thead>
<tr>
<th></th>
<th>Agenda</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RECOMMENDATION: VIOLENCE AND GRAPHIC CONTENT INTERSTITIALS</td>
</tr>
<tr>
<td>2</td>
<td>HEADS UP: VOTER AND CENSUS INTERFERENCE</td>
</tr>
</tbody>
</table>
RECOMMENDATION: Violence and Graphic Content Interstitials
Violence and Graphic Content Interstitials

Overview

Issue: We want to re-consider our approach to application of interstitials (i.e. labels) to violence and graphic content. In doing so, however, we want to be careful that we aren’t discouraging efforts to raise awareness about issues, which may be the intention of sharing graphic imagery.

Summary to Date:
• Consulted with 18 external stakeholders
• Convened 4 working group meetings
• Evaluated thousands of pieces of content across several countries to understand how different policy proposals would affect speech

Recommendation: Maintain status quo, further define newsworthy scenarios, and prioritize ongoing initiatives with product, operations and research.
Violence and Graphic Content Interstitials

Status Quo

Remove:
- Celebrating imagery of violence committed against real humans or animals
- Videos showing mutilation of dying, injured, or dead human beings
- Non-sexual child abuse imagery*

Interstitials:
- Photos of dead or injured humans showing dismemberment, visible innards, or burned bodies
- Videos of dismemberment in a medical setting
- Videos of burning people in cremation or self-immolation context
- Imagery of human torture and animal abuse (unless shared to celebrate or glorify, as noted above)
- Imagery of non-medical foreign objects inserted in humans or animals

Leave Up:
- Fictional depiction of graphic violence
- Animal dismemberment in food context or animal fights in the wild depicting visible innards
- Imagery of human organs outside of a human body
- Taxidermy depictions

*Sexual child abuse imagery is strictly prohibited and dealt with under our policies against Child Nudity and Sexual Exploitation of Children.
WARNING:
Graphic Content
Under our status quo policy, the first two examples would be allowed on the platform without an interstitial — the first image shows an animal that has been prepared as part of a meal, and the second is a still from a video depicting policy brutality. The third image, furthest to the right, features body modification as part of a cultural celebration; the image is allowed on Facebook, but we place it behind a label that reads “Mark as Sensitive.”
Violence and Graphic Content Interstitials

Data Analysis

**Goal:** Assess the type of content people are reporting as Violence & Graphic Content across countries and languages that may not be captured under our existing policies

**Findings:**

- Much of the content reported was not violent or graphic in nature
- Content that was violent or graphic tended to feature: animal abuse by humans, animals that were being served as part of a meal, and police brutality
Violence and Graphic Content Interstitials
Policy Research Key Findings

• A wide body of evidence suggests that exposure to graphic violence has significant impacts, even if in a fictional context.

• People are okay with warning screens but they must be precise in their wording so people can make an informed decision about whether to click through and view the content.

• Warning screens suggest Facebook has ‘reviewed’ content and may inadvertently add credibility to some content.

Graphic Violence Interstitials
External Outreach

We spoke to 18 experts globally, including journalists, human rights organizations, animal rights advocates, psychologists, and safety experts.
Graphic Violence Interstitials

External Outreach

Decrease use of interstitials

Status quo + Newsworthiness Framework

Increase use of interstitials
Violence and Graphic Content Interstitials

Option 1 (Recommendation for discussion): Status Quo Policy + Update newsworthy guidelines

- Status quo policy on what we leave up and what we interstitial is laid out on Slide 6
- Update newsworthy guidelines so that we have a more consistent approach to use of interstitials in the case that something is deemed newsworthy

**Pros:**
- Reflects external feedback
- Balances voice and safety
- Empower users to self-regulate
- Enables use of platform for journalism/activism and awareness raising

**Cons:**
- Self-interstitial feature could be perceived as offloading our responsibility onto the community
Violence and Graphic Content Interstitials

Newsworthy Policy Framework

**Add signals** to our newsworthy guidelines that speak to when we should allow content as newsworthy, but place it behind a label

**Signals we already account for to establish newsworthiness:**
- Videos depicting dismemberment, visible innards, charred, or burning bodies of human beings:
  - shared to condemn a terrorist act or organization, documenting an armed combat
  - documenting genocide, war crimes, or crimes against humanity
- Imagery depicting graphic violence when shared to discuss or raise awareness about or condemn a historical act of violence

**Additional signals to consider:**
- Charred/burning people in context of self-immolation
- Sharing in conflict zones
- Accidental events and natural disasters
- Re-victimization signal to be aware of where real PDITI’s are featured in content
Violence and Graphic Content Interstitials

Option 1: Examples

Under option 1, which maintains status quo and involves an update to our newsworthy guidelines, we would treat the previously discussed examples the same, allowing the first two and marking the third as sensitive.
Violence and Graphic Content Interstitials

**Option 2: Increase use of interstitials**

- Increase use of “Mark as Disturbing” and “Mark as Sensitive” interstitials. Example scenarios include: imagery of non-medical foreign objects voluntarily inserted in humans, videos of children immersed in water in religious/rituals context

**Pros:**
- Covers more content perceived to be gross or graphic

**Cons:**
- May limit visibility of content posted by journalists/activists
- Perceived as culturally insensitive
- More interstitials on people’s timelines may create bad user experience
- Could add credibility to misinformation
- Reduced engagement and reach
- Perception of increased censorship
Under Option 2, the Left and Right examples would be marked as sensitive (no-age gate) while the middle example of police brutality would be marked as disturbing (age-gated).
Violence and Graphic Content Interstitials

Next Steps

• Further develop newsworthy guidelines for permissible shares of violent and graphic content

• Work with product teams to build option to self MAD/MAS feature

• Continue supporting automatic application of MAD interstitials

• Work with research teams to further understand user sentiment tied to use of interstitials
APPENDIX
WARNING:  
Graphic Content
Violence against animals

Appendix
Violence/Graphic content with humans

Appendix
Violence/Graphic content with humans

Status quo: Ignore Human organs outside of body
HEADS-UP:
Voter and Census Interference
Voter and Census Interference

Overview

Issue: We prohibit specific types of voter interference, including the misrepresentation of dates, times, and methods of voting. We recognize that trends in civic interference shift and we want to broaden our policies to protect against more types of interference in voting and census contexts. However, expanding our current voter interference policy may limit political speech and pose operational challenges due to variations in electoral systems across the globe.

Goals:
- Expand policies to address census interference and properly balance malicious civic interference against free speech.
Voter and Census Interference

Status Quo – Voter Suppression and Voter Fraud

Remove

- Voter fraud, defined as, any offers to buy or sell votes with cash or gifts.
- Voter suppression, defined as:
  - Misrepresentation of the dates, locations, times, and methods for voting or voter registration.
  - Misrepresentation of who can vote, qualifications for voting, whether a vote will be counted, and what information and/or materials must be provided in order to vote.
Voter and Census Interference

Next Steps

- Conduct policy research on how voting and census interference manifest on the platform
- Engage with external experts'
- Schedule internal working groups
- Engage with cross-functional internal teams to ensure that expansion of the policy aligns with other workstreams related to voter interference
facebook