

POLICY FORUM

FACEBOOK

June 15th, 2021

Recommendation: Brigading & Mass Harassment

Actor & Behavioral Policy | Security Policy

Issue Statement:

Large groups of people targeting communication to one person can be used as a tool for activism and civic engagement. However, it is also used as a tactic to harass or potentially harm people. We are trying to develop a policy framework that protects people from mass harassment and intimidation without penalizing people engaging in legitimate forms of communication.

Sources: XFN feedback; external stakeholder feedback; research; investigations; escalations

Working Groups: 12

External Engagements: 97

Brigading & Mass Harassment

Goals for this Policy Forum

Goal 1

Ratify an **impact-based policy framework** that captures the range of brigading and mass harassment behaviors on our platforms.

Goal 2

Approve recommendation of a **slate of initial policies** to begin enforcement on two types of brigading:

- (a) Network-level takedowns for highly adversarial brigading
- (b) Actor-level penalties for high-impact mass harassment

Brigading & Mass Harassment

Research & External Engagement

Research

Key Points:

- While brigading and mass harassment are well defined as concepts, we lack a broadly accepted method of detecting or measuring brigading when it occurs.
- Research suggests that many people think coordinated brigading should not be allowed in certain contexts, but some groups are relatively more permissive.

External Engagement

Key Points:

- Stakeholders agreed that coordinated attacks could have more far-reaching consequences on victims than organic harassment. However, some pointed out that it would be challenging for FB to differentiate between the two - networked behavior does not exclude organic behavior and vice versa.
- While a majority agreed that it would be appropriate to give vulnerable individuals greater protections from brigading, there were differences in stakeholders' interpretations of what constitutes a vulnerable individual.

Brigading & Mass Harassment

Goals for this Policy Forum

Goal 1

Ratify an **impact-based policy framework** that captures the range of brigading and mass harassment behaviors on our platforms.

Goal 2

Approve recommendation of a **slate of initial policies** to begin enforcement on two types of brigading:

- (a) Network-level takedowns for highly adversarial brigading
- (b) Actor-level penalties for high-impact mass harassment

Brigading & Mass Harassment

Policy Framework

IMPACT TO TARGET

BRIGADING ELEMENTS

	Low Impact	Medium Impact	High Impact
Actor	Private individuals, political constituents	Public figures, high reach accounts	State-linked actors, adversarial networks
Coordination (on- or off-platform)	Uncoordinated: Organic activity, no evidence of any coordination	Directed coordination: Calls to action, encouragement	Mobilized coordination: Evidence of extensive, covert, and/or adversarial networks
Behavior	Non-violating content or behavior on public surfaces	Objectionable content or behavior on private surfaces, including inboxes	Violating content or behavior on public or private surfaces or targeting based on protected characteristics
Target	Corporations, brands, governments, entities	Anyone targeted as an individual	Individuals vulnerable to brigading (e.g. minors, human rights defenders)

Brigading & Mass Harassment

Goals for this Policy Forum

Goal 1

Ratify an **impact-based policy framework** that captures the range of brigading and mass harassment behaviors on our platforms.

Goal 2

Approve recommendation of a **slate of initial policies** to begin enforcement on two types of brigading:

- (a) Network-level takedowns for highly adversarial brigading
- (b) Actor-level penalties for high-impact mass harassment

Brigading & Mass Harassment

Goals for this Policy Forum

Goal 1

Ratify an **impact-based policy framework** that captures the range of brigading and mass harassment behaviors on our platforms.

Goal 2

Approve recommendation of a **slate of initial policies** to begin enforcement on two types of brigading:

(a) Network-level takedowns for highly adversarial brigading

(b) Actor-level penalties for high-impact mass harassment

Brigading & Mass Harassment

Recommended Policies: Network-Level

POLICY INTERVENTIONS

	1a/ Network + Account Disables for State-Linked Brigading	1b/ Networks + Account Disables for Highly Adversarial Networks
Actor	States and state-affiliated entities	Adversarial networks
Coordination	Mobilized coordination	Mobilized coordination
Behavior	Any	Objectionable content or behavior on public or private surfaces
Target	Anyone targeted as an individual	Anyone targeted as an individual

**Note: Enforcements for network takedowns will require approval through Disruption Reviews.*

Brigading & Mass Harassment

Goals for this Policy Forum

Goal 1

Ratify an **impact-based policy framework** that captures the range of brigading and mass harassment behaviors on our platforms.

Goal 2

Approve recommendation of a **slate of initial policies** to begin enforcement on two types of brigading:

(a) Network-level takedowns for highly adversarial brigading

(b) Actor-level penalties for high-impact mass harassment

Brigading & Mass Harassment

Recommended Policies: Actor-Level

POLICY INTERVENTIONS

		2/ Mass Harassment of Individuals Vulnerable to Brigading	3/ Mass Harassment on Private Surfaces
POLICY FRAMEWORK	Actor	Any	Any
	Coordination	Evidence of directed coordination on- or off-platform	Evidence of directed coordination on- or off-platform
	Behavior	Any content or behavior on public or private surfaces	Objectionable content or behavior on private surfaces, including inboxes
	Target	Individuals vulnerable to brigading	Anyone targeted as an individual

Brigading & Mass Harassment

Recommended Policies: Actor-Level

POLICY INTERVENTIONS

POLICY FRAMEWORK

	2/ Mass Harassment of Individuals Vulnerable to Brigading	3/ Mass Harassment on Private Surfaces
Actor	Any	Any
Coordination	Evidence of directed coordination on- or off-platform	Evidence of directed coordination on- or off-platform
Behavior	Any content or behavior on public or private surfaces	Objectionable content or behavior on private surfaces, including inboxes
Target	Individuals vulnerable to brigading	Anyone targeted as an individual
Enforcement	Remove content & issue strike	Messenger feature blocks