

Hateful Conduct

 help.x.com/en/rules-and-policies/hateful-conduct-policy

July 17, 2024

Overview

April 2023

You may not directly attack other people on the basis of race, ethnicity, national origin, caste, sexual orientation, gender, gender identity, religious affiliation, age, disability, or serious disease.

X's mission is to give everyone the power to create and share ideas and information, and to express their opinions and beliefs without barriers. Free expression is a human right – we believe that everyone has a voice, and the right to use it. Our role is to serve the public conversation, which requires representation of a diverse range of perspectives.

We recognize that if people experience abuse on X, it can jeopardize their ability to express themselves. Research has shown that some groups of people are disproportionately targeted with abuse online. For those who identify with multiple underrepresented groups, abuse may be more common, more severe in nature, and more harmful.

We are committed to combating abuse motivated by hatred, prejudice or intolerance, particularly abuse that seeks to silence the voices of those who have been historically marginalized. For this reason, we prohibit behavior that targets individuals or groups with abuse based on their perceived membership in a protected category.

If you see something on X that you believe violates this policy, please [report it to us](#).

What is in violation of this policy?

We will review and take action against reports of accounts targeting an individual or group of people with any of the following behavior, whether within Posts or Direct Messages.

Hateful references We prohibit targeting individuals or groups with content that references forms of violence or violent events where a protected category was the primary target or victims, where the intent is to harass. This includes, but is not limited to media or text that refers to or depicts:

- genocides, (e.g., the Holocaust);
- lynchings.

Incitement

We prohibit inciting behavior that targets individuals or groups of people belonging to protected categories. This includes:

- inciting fear or spreading fearful stereotypes about a protected category, including asserting that members of a protected category are more likely to take part in dangerous or illegal activities, e.g., “all [religious group] are terrorists.”
- inciting others to harass members of a protected category on or off platform, e.g., “I’m sick of these [religious group] thinking they are better than us, if any of you see someone wearing a [religious symbol of the religious group], grab it off them and post pics!”
- inciting others to discriminate in the form of denial of support to the economic enterprise of an individual or group because of their perceived membership in a protected category, e.g., “If you go to a [religious group] store, you are supporting those [slur], let’s stop giving our money to these [religious slur].” This may not include content intended as political in nature, such as political commentary or content relating to boycotts or protests.

Note: content intended to incite violence against a protected category is prohibited under [Violent Speech](#).

Slurs and Tropes We prohibit targeting others with repeated slurs, tropes or other content that intends to degrade or reinforce negative or harmful stereotypes about a protected category. In some cases, such as (but not limited to) severe, repetitive usage of slurs, or racist/sexist tropes where the context is to harass or intimidate others, we may require Post removal. In other cases, such as (but not limited to) moderate, isolated usage where the context is to harass or intimidate others, we may limit Post visibility as further described below.

Dehumanization

We prohibit the dehumanization of a group of people based on their religion, caste, age, disability, serious disease, national origin, race, ethnicity, gender, gender identity, or sexual orientation.

Hateful Imagery

We consider hateful imagery to be logos, symbols, or images whose purpose is to promote hostility and malice against others based on their race, religion, disability, sexual orientation, gender identity or ethnicity/national origin. Some examples of hateful imagery include, but are not limited to:

- symbols historically associated with hate groups, e.g., the Nazi swastika;
- images depicting others as less than human, or altered to include hateful symbols, e.g., altering images of individuals to include animalistic features; or
- images altered to include hateful symbols or references to a mass murder that targeted a protected category, e.g., manipulating images of individuals to include yellow Star of David badges, in reference to the Holocaust.

Media depicting hateful imagery is not permitted within live video, account bio, profile or header images. All other instances must be marked as sensitive media. Additionally, sending an individual unsolicited hateful imagery is a violation of this policy.

Hateful Profile

You may not use hateful images or symbols in your profile image or profile header. You also may not use your username, display name, or profile bio to engage in abusive behavior, such as targeted harassment or expressing hate towards a person, group, or protected category.

Do I need to be the target of this content for it to be a violation of the X Rules?

Some Posts may appear to be hateful when viewed in isolation, but may not be when viewed in the context of a larger conversation. For example, members of a protected category may refer to each other using terms that are typically considered as slurs. When used consensually, the context behind these terms is not abusive, but a means to reclaim terms that were historically used to demean individuals.

When we review this type of content, it may not be clear whether the context is to abuse an individual on the basis of their protected status, or if it is part of a consensual conversation. To help our teams understand the context, we sometimes need to hear directly from the person being targeted to ensure that we have the information needed prior to taking any enforcement action.

Note: individuals do not need to be a member of a specific protected category for us to take action. We will never ask people to prove or disprove membership in any protected category and we will not investigate this information.

What happens if you violate this policy?

Under this policy, we take action against behavior that targets individuals or an entire protected category with hateful conduct, as described above. Targeting can happen in a number of ways, for example, mentions, including a photo of an individual, referring to someone by their full name, etc.

When determining the penalty for violating this policy, we consider a number of factors including, but not limited to the severity of the violation and an individual's previous record of rule violations. The following is a list of potential enforcement options for content that violates this policy:

- Making content less visible on X by:
 - Removing the Post from search results, in-product recommendations, trends, notifications, and home timelines
 - Restricting the Post discoverability to the author's profile
 - Downranking the Post in replies
 - Restricting Likes, replies, Reposts, Quote, bookmarks, share, pin to profile, or engagement counts
 - Excluding the Post from having ads adjacent to it
- Excluding Posts and/or accounts in email or in-product recommendations.
- Requiring Post removal.
For example, we may ask someone to remove the violating content and serve a period of time in read-only mode before they can Post again.
- Suspending accounts that violate our Hateful Profile policy.

Learn more about our [range of enforcement options](#).

If someone believes their account was suspended in error, they can [submit an appeal](#).

Share this article
