

Hateful conduct policy



1. [Help Center](https://help.twitter.com/) ^
2. [Safety and cybercrime](https://help.twitter.com/en/rules-and-policies#safety-and-cybercrime) ^

Hateful conduct policy

Hateful conduct: <https://help.twitter.com/rules-and-policies/twitter-rules#hateful-conduct> You may not promote violence against or directly attack or threaten other people on the basis of race, ethnicity, national origin, caste, sexual orientation, gender, gender identity, religious affiliation, age, disability, or serious disease. We also do not allow accounts whose primary purpose is inciting harm towards others on the basis of these categories.

Hateful imagery and display names: 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.

Rationale

Twitter'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 Twitter, it can jeopardize their ability to express themselves. Research has shown that some groups of people are disproportionately targeted with abuse online. This includes; women, people of color, lesbian, gay, bisexual, transgender, queer, intersex, asexual individuals, marginalized and historically underrepresented communities. 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 Twitter that you believe violates our hateful conduct policy, please [report it to us](#).

When this applies

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 Tweets or Direct Messages.

Violent threats

We prohibit content that makes violent threats against an identifiable target. Violent threats are declarative statements of intent to inflict injuries that would result in serious and lasting bodily harm, where an individual could die or be significantly injured, e.g., “I will kill you.”

Note: we have a zero tolerance policy against violent threats. Those deemed to be sharing violent threats will face immediate and permanent suspension of their account.

Wishing, hoping or calling for serious harm on a person or group of people

We prohibit content that wishes, hopes, promotes, incites, or expresses a desire for death, serious bodily harm, or serious disease against an entire protected category and/or individuals who may be members of that category. This includes,

but is not limited to:

- Hoping that an entire protected category and/or individuals who may be members of that category dies as a result of a serious disease, e.g., “I hope all [nationality] get COVID and die.”
- Wishing for someone to fall victim to a serious accident, e.g., “I wish that you would get run over by a car next time you run your mouth.”
- Saying that a group of individuals deserve serious physical injury, e.g., “If this group of [slur] don’t shut up, they deserve to be shot.”
- Encouraging others to commit violence against an individual or a group based on their perceived membership in a protected category, e.g., “I’m in the mood to punch a [racial slur], who’s with me?”

References to mass murder, violent events, or specific means of violence where protected groups have been the primary targets or victims

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 against protected categories

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

- to incite fear or spread 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.”
- to incite 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!”
- to incite 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 that content intended to incite violence against a protected category is prohibited under **Wishing, hoping, or calling for serious harm on a person or groups of people.**

We prohibit targeting individuals and groups with content intended to incite fear or spread 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.”

Repeated and/or non-consensual slurs, epithets, racist and sexist tropes, or other content that degrades someone

We prohibit targeting others with repeated slurs, tropes or other content that intends to dehumanize, degrade or reinforce negative or harmful stereotypes about a protected category. This includes targeted misgendering or deadnaming of transgender individuals. We also 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. In some cases, such as (but not limited to) severe, repetitive usage of slurs, epithets, or racist/sexist tropes where the primary intent is to harass or intimidate others, we may require Tweet removal. In other cases, such as (but not limited to) moderate, isolated usage where the primary intent is to harass or intimidate others, we may limit Tweet visibility as further described below.

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 our [abusive behavior policy](#).

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

Some Tweets 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 intent 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 intention 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.

Consequences

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:

- Downranking Tweets in replies, except when the user follows the Tweet author.
- Making Tweets ineligible for amplification in Top search results and/or on timelines for users who don't follow the Tweet author.
- Excluding Tweets and/or accounts in email or in-product recommendations.
- Requiring Tweet 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 Tweet again. Subsequent violations will lead to longer read-only periods and may eventually result in permanent suspension.
- Suspending accounts whose primary use we've determined is to engage in hateful conduct as defined in this policy, or who have shared violent threats.

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

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

[appeal](https://help.twitter.com/forms/general?subtopic=suspended) (https://help.twitter.com/forms/general?subtopic=suspended).

Share this article



Was this article helpful?^{*}



Submit



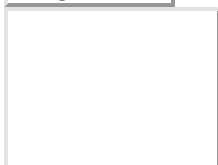
© 2022 Twitter, Inc.

[Cookies](https://help.twitter.com/rules-and-policies/twitter-cookies) (https://help.twitter.com/rules-and-policies/twitter-cookies)

[Privacy](https://twitter.com/privacy) (https://twitter.com/privacy)

[Terms and conditions](https://twitter.com/tos) (https://twitter.com/tos)

English



[Help Center](https://help.twitter.com/) (https://help.twitter.com/)