

Singapore Code of Practice for Online Safety Transparency Report

Submitted: July 2024

Re-submitted with clarifications: December 2024

Code Obligations:

Paragraph 8: End-users must be able to use the Service in a safe manner. In this regard, the Service must put in place measures to minimize end-users' exposure to harmful content, empower end-users to manage their safety on the Service and mitigate the impact on end-users that may arise from the propagation of harmful content.

And

Paragraph 9: Children in particular, may lack the capacity or experience to deal with the information and content available online and will need more protection to ensure a safer online space for them. In this regard, the Service must therefore also have specific measures to protect children from harmful content.

Section A: User Safety

Please provide information on the measures in place for all End-users in Singapore:

YouTube's systems are designed to connect users with high quality content from authoritative sources and minimise exposure to harmful content. In addition, YouTube has various policies which set what is not allowed on YouTube. These policies can be accessed via this <u>landing page</u> in YouTube's Help Centre. For example, YouTube's <u>Community Guidelines</u> include policies that cover the categories of harmful content in Paragraph 4 of the Code of Practice.

YouTube's approach to curbing the spread of harmful content involves removing content that violates its policies as quickly as possible, raising high quality information in rankings and recommendations, and rewarding trusted eligible creators and artists. YouTube applies these principles globally. YouTube uses a combination of people and machine learning to detect harmful content automatically and at scale. Machine learning is well-suited to detect patterns, which helps YouTube find content similar to other content that YouTube has already removed, even before it is viewed.

YouTube also recognises the best way to quickly remove content is to anticipate problems before they emerge. YouTube's Intelligence Desk, a team of specialised analysts, identifies potentially violative trends—whether new vectors of misinformation or dangerous internet challenges—and the risks they pose. They also monitor ongoing threats like extremist conspiracy theories, both tracking their prevalence across social media and evaluating how they morph over time.

In this report, we highlight additional details about specific measures in places to minimise users' exposure to harmful content, empower users to manage their safety on YouTube and mitigate the impact on users that may arise from the propagation of harmful content.

Please provide information on any differentiated or additional measures in place for Children:

While we provide kids and teenagers with a space and tools to navigate their imagination and curiosity, we also understand that these same young people need special protections and considerations when it comes to what they experience and discover online. Accordingly, we give parents choices to guide their kids' YouTube journey and invest heavily in the technology and teams that help provide kids and families with the best protection possible. These are:

YouTube Kids

<u>YouTube Kids</u> is a separate app built from the ground up to be a safer and simpler experience for kids to explore, with tools for parents and caregivers to guide their journey. This separate app is a filtered version of YouTube and has a much smaller set of channels and videos available than the YouTube app and website.

On YouTube Kids, we work to identify content that is age-appropriate, adheres to our quality principles, and is diverse enough to meet the varied interests of kids globally. This includes specific content policies designed with feedback from our Youth and Families Advisory Committee, which is made up of external specialists in children's media, child development, digital learning, and citizenship. These policies decide what kind of content is eligible for each content setting (Pre-school, Younger and Older content, as described here). When parents choose a content setting for their child, they will see content eligible for that setting. The setting parents choose will impact what their children find in searches or in recommended videos. Our automated systems and human reviewers work to identify suitable content based on the setting they have selected.

In addition to content settings, parents can tailor the YouTube Kids experience for their child by adjusting specific parental controls and settings. These controls and settings let parents (please note this is not an exhaustive list):

- Block or share content with their child
- Disable autoplay
- Limit screen time on YouTube Kids

Supervised Experiences

A <u>supervised experience</u> on YouTube is a parent-managed version of regular YouTube for children under 13. With a supervised account, parents select a <u>content setting</u> that limits the videos and music children under 13 can find and play. <u>Supervised accounts</u> also limit the features they can use, the default account settings, and the ads they see.

As of March 2024, Supervised Experiences were not available for parents or legal guardians on accounts created after a child turns 13. *

However, in August 2024, <u>Google announced</u> parents and teens will be able to link their accounts, enabling parents to gain visibility into their teens' activity on YouTube. *

Based on YouTube's <u>Terms of Service</u>, YouTube users must be at least 13 years old to use YouTube unless they are enabled by their parent or a legal guardian. YouTube is committed to ensuring as safe as possible an experience for our young users, such as in the following ways:

Made for Kids Quality Principles

We have developed a set of <u>quality principles</u> to help guide creators who upload kids and family content. These principles were developed with child development specialists, based on extensive research, and supplement our Community Guidelines, which help create a safe viewing experience for everyone. The quality principles for kids and family content may affect a channel's performance. High-quality content (e.g. content about being a good person, learning and inspiring curiosity, or creativity, play and a sense of imagination) gets raised up in recommendations. The principles also guide decisions both for inclusion in YouTube Kids and channel and video monetisation. On the other hand, content that demonstrates low-quality (e.g. content that encourages negative behaviours or attitudes, or sensational or misleading content) is reduced in recommendations. Channels found to have a strong focus on low-quality "made for kids" content may be suspended from the YouTube Partner Program. If an individual video is found to violate these quality principles, it may also get limited or no ads.

Age restriction and safeguards

Google uses a combination of people and machine learning tools to detect and flag potentially age inappropriate content for minors, which is then age-restricted. YouTube also provides an option for creators to age-restrict their own videos. This type of age restriction is self-imposed and is not a result of a review by YouTube. When a video is age-restricted, viewers must be signed in and 18 years of age or older to view it. These videos are not shown in certain sections of YouTube. Age-restricted videos may also have limited or no ads. Videos that creators proactively age restrict are still subject to YouTube's Community Guidelines. If YouTube decides that a video should be age restricted, a permanent age restriction will be applied. This restriction happens even if the creator also age restricts the video.

Additionally, see our response to Paragraph 18 below for more information about YouTube's <u>safeguards limiting repeated recommendations</u> of certain content to support the well-being and mental health of young people.

Child safety policy

To better protect minors, YouTube may disable certain features at both the channel and content level if children are featured, pursuant to its <u>child safety policy</u>. YouTube also disallows children under the age of 13 to appear in a live stream unless they are visibly accompanied by an adult. Channels not in compliance with this policy may temporarily or permanently lose their ability to live stream.

In this report, we highlight additional details about specific measures in place to protect children from harmful content.

* Additional information added in December 2024

Paragraph 11: End-users' exposure to harmful content must be minimised via reasonable and proportionate measures. These measures include, but are not limited to, a set of community guidelines and standards, and content moderation measures that are put in place and effected by the Service. The Service's community guidelines and standards must address the categories of harmful content in paragraph 4 and must be published.

Please provide information on the community guidelines and standards in place and how they address the categories of harmful content in paragraph 4.

Please provide information on the content moderation measures in place and how they are enforced or effected. Please include screenshots or evidence where possible.

Community Guidelines on YouTube

YouTube enforces a broad range of policies to help build a safe community. These policies include, but are not limited to, YouTube's <u>Community Guidelines</u>, which broadly cover the categories of harmful content in Paragraph 4 of the Code of Practice. YouTube applies these policies globally.

These include the following:

- Nudity and sexual content policy
- Violent or graphic content policy
- Harmful or dangerous content policy
- Suicide and self-harm policy
- Harassment and cyberbullying policy
- <u>Misinformation policy</u>, including YouTube's <u>medical misinformation policy</u>
- Illegal or regulated goods or services policies
- Violent extremist or criminal organisations policy

Content Moderation on YouTube

YouTube continues to invest in automated detection systems, and relies on both human evaluators and machine learning to detect and enforce policies against harmful content. The vast majority of content reviewed and enforced on YouTube is first detected by automated systems. However, after potentially violative content has been detected by automated systems, content moderators may review the content to confirm the decision.

Human reviewers or content moderators also play a key role in content moderation at YouTube. Although technology has become very helpful in identifying some kind of harmful content, humans are able to apply a more nuanced approach to assessing content. In order to safeguard against content actions that could potentially continue to or exacerbate adverse impacts due to allowing or removing content, YouTube utilises international human rights standards to guide policy and enforcement decision-making, considering how content could adversely impact the rights of an individual, community, or society as a whole, or further understanding of social, political, cultural, civic, and economic affairs. As an example of public-interest informed content moderation, YouTube carves out exceptions to enforcement guidelines for material that is Educational Documentary, Scientific, and/or Artistic (EDSA). Content that falls under those exceptions are crucial to understanding the world and to chronicling history, whether it is documenting wars and revolutions, or artistic expression that may include nudity. Consequently, YouTube takes great care in helping reviewers understand the EDSA exceptions when reviewing content.

Paragraph 12: End users must have access to tools that enable them to manage their own safety and effectively minimise their exposure to, and mitigate the impact of, harmful content and unwanted interactions on the Service. Such tools may include:

- (a): Tools to restrict visibility of harmful content and/or unwanted comments
- (b): Tools to limit visibility of the end-user's account, including profile and content, as well as contact and/or interactions with other end-users
- (c): Tools to limit location sharing

Please provide information on the following, and include screenshots or evidence where possible.

All End-users in Singapore:

Paragraph 12(a): Tools to restrict visibility of harmful content and/or unwanted comments

Tools to report abuse or inappropriate content on YouTube

For content believed to violate YouTube's <u>Community Guidelines</u>, users can report content using YouTube's <u>flagging feature</u>, which is available to signed-in users in Singapore via computer (desktop or laptop), mobile devices, and other surfaces. Details on how to report different types of content using YouTube's flagging feature is outlined in YouTube's <u>Help Centre</u>. Users can also submit a <u>legal complaint</u> if content identified may violate local law.

Restricted Mode on YouTube

<u>Restricted Mode</u> is an optional setting that users can use on YouTube. This feature can help screen out potentially mature content that you or others using your devices may prefer not to view.

Managing recommendations and search results

Users can manage the recommendations they get by choosing to view, delete, or turn on or off their YouTube watch and search history at any time. Specifically, if users do not want to see recommendations at all, they can turn off and clear their YouTube watch history. For users with YouTube watch history off and no significant prior watch history, the homepage will show the search bar and the Guide menu, with no feed of recommended videos. This allows users to take steps to minimise the chances of them getting recommendations for content that they feel is harmful. Finally, users can also

use the "Not interested" and "Don't recommend channel" options to prohibit specific videos and channels from showing up in their recommendations.

Moderating comments and live chats on YouTube

YouTube offers users who upload videos the opportunity to control various <u>settings</u> to restrict the visibility of comments on videos that they upload. Creators can choose to turn comments on, pause them, or turn them off for a specific video. YouTube also offers the ability for creators to <u>hide comments from specific users</u> from the rest of the creator's audience, which means the hidden channel's comments won't show on any videos on the creator's channel, and they cannot create clips from the creator's videos or live streams. When comments are turned on, users are given the option to decide which ones, if any, are held for review. Comments held for review are available in over 100 languages and are not publicly visible unless a user approves them. Viewers of YouTube videos are also able to <u>block messages on live chat</u>, if they no longer want to see that viewer's messages on live chat.

Users can also select moderators to help review and manage comments other users leave on their videos or the messages participants send during a live chat. There are two types of moderators: a standard moderator and managing moderator. A standard moderator can remove any inappropriate or potentially abusive or offensive content or hide another user from viewing a channel. A managing moderator has the same privileges as a standard moderator, in addition to being able to enable features to detect spam or self-promotion in comments, turn live chat on or off, limit how often other users can send a chat message, and block live chat messages that contain or closely match certain words.

All End-users in Singapore:

Paragraph 12(b): Tools to limit visibility of the end-user's account, including profile and content, as well as contact and/ or interactions with other end-users

Users who wish to limit visibility of content uploaded to their channel or interaction with users can adjust channel-level settings, comment settings, or privacy settings.

Limit channel visibility

Users can <u>hide</u> their channel to make their channel name, videos, likes, subscriptions, and subscribers private. If a user wishes to make their content viewable to others, or if they want to upload, comment, or use playlists, they will need to re-enable their channel following the steps outlined <u>here</u>.

Limit contact with other users

As mentioned in our response to Paragraph 12(a), YouTube offers users who upload videos the ability to control their <u>comment settings</u>, including pausing or turning off comments for a specific video, and hiding comments from a specific user. Additionally, users can add a list of words and phrases that they do not want to show in comments to "blocked words" lists.

Manage privacy settings on content

By adjusting privacy settings, users can limit visibility into their <u>videos</u> and <u>playlists</u>. Both videos and playlists can be set to private or unlisted. Private videos and playlists can only be seen by the user and whomever they choose, while unlisted videos and playlists can be seen and shared by anyone with the link to the content. Private and unlisted

videos won't appear in the Videos tab of a channel homepage or show up in YouTube's search results. Unlisted videos may show up in search results if added to another user's public playlist.

All End-users in Singapore:

Paragraph 12(c): Tools to limit location sharing

YouTube does not by default display the location of creators in their respective videos, and also does not display the location of users on any comments.

Paragraph 13: End-users must be able to easily access information related to online safety on the Service. Such information must be easy to understand and must include the availability of tools and local information, including Singapore-based safety resources or support centres, if available. The service should seek to implement, support and/ or maintain programmes and initiatives to educate and raise awareness of such information.

And

Paragraph 21: Children must be able to easily access information related to online safety on the Service. Such information must be easily understood by children and must include information on tools available to protect children harmful and/or inappropriate content and unwanted interactions, as well as local information, including Singapore-based safety resources or support centres, if available. The Service should seek to implement, support and/or maintain programmes and initiatives to educate and raise awareness of such information.

Please provide information on the measures, and include screenshots or evidence where possible.

All End-users in Singapore:

YouTube has published a Help Centre article on <u>Information on online safety</u> specifically for users in Singapore. This article outlines information about:

- Policies in place to address harmful content
- How to report Community Guidelines-violative or illegal content
- Locally based organisations in Singapore to assist users in need of suicide and self-harm support
- Parental controls and settings for YouTube Kids profiles & supervised accounts

In addition to the <u>YouTube Help Centre</u>, YouTube regularly updates its <u>YouTube for Families Help Centre</u> to share safety resources for families with children or teens who use YouTube Kids or Supervised Experiences. This Help Centre page is accessible to everyone and provides information about YouTube's products and features that parents and caregivers can use to guide their kids' YouTube journey and choose the best YouTube experience for their family.

For teens, YouTube has partnered with experts like the World Health Organization (WHO) and Common Sense Networks to develop public industry-wide resources (example here) related to teens and online well-being. YouTube also has a family guide for supervised experiences that outlines tips and tricks that parents and caregivers can leverage to help build healthy digital habits.

For creators, YouTube makes safety resources available in the <u>Creator Safety Centre</u>. The resources cover a range of topics including, but not limited to: growing a channel safely, managing unwanted behavior from others, and finding support for matters related to physical danger and violent threads as well as depression, suicide or thoughts of self-harm.

Please provide information on the measures, and include screenshots or evidence where possible.

Children:

See response to Paragraph 13, which also applies to users who are children.

Paragraph 14: End-users who use high-risk search terms such as, but not limited to, terms relating to self-harm and suicide on the Service must be actively offered relevant safety information (stated in paragraph 13) such as, but not limited to, local suicide prevention hotlines, if available.

And

Paragraph 22: Children who use high-risk search terms, such as, but not limited to, terms relating to self-harm and suicide, on the Service must be actively offered relevant safety information (stated in paragraph 21) such as, but not limited to, local suicide prevention hotlines, if available.

Please provide information on the measures, and include screenshots or evidence where possible.

All End-users in Singapore:

Protective measures are in place for high-risk search terms used by any user (including a user that is a child). When a user in Singapore searches for content related to self-harm and suicide, search results will display a <u>crisis resource panel</u> encouraging them to seek support and share contact information of locally based organisations like <u>Samaritans of Singapore</u>.

YouTube <u>recently expanded</u> crisis resource panels into a new full-page experience that better helps viewers pause for a moment. This full-page experience allows viewers to more prominently see resources for the third-party crisis hotlines run by locally based organisations.

An example of the full-page experience of the crisis resource panel in Singapore is shown in the screenshot below:

suicide	X Q .
	i i
	You're not alone
	If you or someone you know is having a hard time, talk to somebody today.
	Samaritans of Singapore
	& Call 1800-221-4444
	Free - Confidential - Hours: Available 24 hours
	The following results may be about suicide or self-harm. Show anyway

Please provide information on the measures, and include screenshots or evidence where possible.

Children:

See response to Paragraph 14, which also applies to users who are children.

Paragraph 15: End-users' exposure to child sexual exploitation and abuse material and terrorism content on the Service must be minimized through the use of technologies and processes. These technologies and processes must proactively detect and swiftly remove child sexual exploitation and abuse material and terrorism content as technically feasible, such that the extent and length of time to which such content is available on the Service is minimised.

Please provide information on the measures, and include screenshots or evidence where possible.

All End-users in Singapore:

YouTube relies on a combination of people and technology to flag inappropriate content and enforce YouTube's Community Guidelines. Flags can come from our automated flagging systems, from members of the Priority Flagger program, or from users in the broader YouTube community. YouTube developed powerful machine learning that detects content that may violate YouTube's policies and sends it for human review. In some cases, that same machine learning automatically takes an action, if there is high confidence that content is violative given information about similar or related content that has been previously removed.

We elaborate further on the specific measures we have in place in relation to CSAM and terrorism content below:

Paragraph 16: End-users must be protected from preparatory child sexual exploitation and abuse activity and terrorism activity on the Service through reasonable and proportionate steps taken by the Service to proactively detect and swiftly remove preparatory child sexual exploitation and abuse activity (such as online grooming for child sexual abuse) and terrorism activity (such as glorifying or endorsing terrorist activities and recruitment).

Minimising exposure to child sexual abuse material (CSAM) and protecting end-users from preparatory child sexual exploitation

YouTube has strict policies and robust operations in place to tackle content and behaviour that is harmful or exploitative to children. Under YouTube's <u>Child Safety</u> policy (as mentioned in our response to Paragraph 9 above and Paragraph 17 below), YouTube prohibits content that puts minors at risk including areas such as unwanted sexualisation and child sexual abuse material (CSAM).

We have heavily invested in engineering resources to detect child sexual abuse material (CSAM) in ways that are precise and effective, and have long used this technology to prevent the distribution of known child sexual abuse imagery (CSAI) videos on YouTube. Our proprietary <u>CSAI Match technology</u> allows us to detect known CSAI videos. In cases where a video contains CSAI or a user solicits CSAI through comments or other communications, our team reports it to the <u>National Center for Missing and Exploited Children (NCMEC)</u>, who then liaise with global law enforcement agencies.

YouTube also deploys safety by design principles to deter users from seeking out child sexual abuse material (CSAM) on YouTube. YouTube blocks search results that lead to child sexual abuse imagery or material that appears to sexually victimise, endanger or otherwise exploit children.

Google's <u>Combating Child Sexual Abuse Material transparency report</u> includes data about YouTube's efforts to detect, remove, and report CSAM and CSAI. Learn about CSAI Match and our other tools to identify CSAM at scale <u>here</u>.

Minimising exposure to terrorist content

Under YouTube's <u>violent extremist or criminal organisations policy</u>, content intended to praise, promote, or aid violent extremist or criminal organisations is not allowed on YouTube. These organisations are not allowed to use YouTube for any purpose, including recruitment.

Additionally, we use hash matching (or "digital fingerprints") technology to detect known violent extremist content. Previously reviewed content determined to be violative of our policies is hashed, and these hashes are compared against other content hosted on our products. When a match is found, the content is reviewed to verify whether it indeed violates our policies. If it does, the content is removed and is used to train our machines for better coverage in the future.

In 2016, YouTube created a <a href="https://mash-sharing.com/hash-sharing.

Paragraph 17: Besides harmful content, children's exposure to inappropriate content must also be minimised through reasonable and proportionate measures. These measures include, but are not limited to, a set of community guidelines and standards and content moderation measures put in place and effected by the Service that are appropriate for children. These community guidelines and standards must minimally address the following categories of content, and must be published:

(a): Sexual content

(b): Violent content

(c): Suicide and self-harm content

(d): Cyberbullying content

Please provide information on the measures, and include screenshots or evidence where possible.

Children:

As mentioned in our response to Paragraph 9, YouTube has a broad range of Community Guidelines and content moderation measures to help limit users' (including children's) exposure to harmful content.

Our Community Guidelines include, but are not limited to, the following:

- Nudity and sexual content policy
- Violent or graphic content policy
- Harmful or dangerous content policy
- Suicide and self-harm policy
- Harassment and cyberbullying policy

In addition to such generally applicable policies, we highlight that YouTube has the following policies to further protect children, in relation to the 4 categories of harmful content mentioned in Paragraph 17 of the Code:

Child Safety Policy

YouTube does not allow content that endangers the emotional and physical well-being of minors, including content that features sexualization of minors, content that targets young minors and families but contain age-inappropriate themes, and content showing

cyberbullying and harassment toward minors. A minor is someone under 18 years old. Examples of content that will be removed if it violates this policy include (please note this is not an exhaustive list):

- Sexually explicit content featuring minors and content that sexually exploits minors.
- Content that targets young minors and families but contains sexual themes, violence, obscene, or other mature themes not suitable for young audiences.
- Content inflicting or advocating for the infliction of physical, sexual or emotional maltreatment or neglect of a child, including inflicting emotional distress on minors.
- Content that intends to shame, deceive or insult a minor, or encourages others to bully or harass.

Age Restriction

Further, YouTube may age-restrict (i.e. such videos will not be viewable to users below 18 years of age or are signed out):

- Nudity and sexually suggestive content
 - A video that invites sexual activity, such as provocative dancing or fondling
 - A video where the subject is in a pose that is intended to sexually arouse the viewer
 - A video where the subject is in clothing that is considered unacceptable in public contexts, such as lingerie
- Violent or graphic content
 - A video with context showing survivor's injuries in a major road accident
 - A video focused on violent or gory imagery, such as focusing solely on the most graphically violent part of a film or video game

Paragraph 18: Children must not be targeted to receive content that the Service is reasonably aware to be detrimental to their physical or mental well-being. Such content includes the categories of harmful and/or inappropriate content in paragraphs 4 and 17. In this regard, content targeting refers, but is not limited to, advertisements, promoted content and content recommendations.

Please provide information on the measures, and include screenshots or evidence where possible.

Children:

First and foremost, as mentioned in our response to Paragraph 9, YouTube seeks to remove content that violates our Community Guidelines. YouTube also has certain policies specific to children, which are outlined in our response to Paragraph 17 above.

In addition, within the YouTube Kids and within Supervised Experiences on the main YouTube platform, parents can select the content setting they want their children to have access to, and from which recommendations are made. In YouTube Kids, parents can choose between 3 content settings as described here. For Supervised Experiences on the main YouTube platform, parents can also select a content setting for their child, as described here.

Parents or guardians can also customise young users' access to content by <u>blocking</u> <u>content for their child's YouTube Kids profile</u> or for their child's supervised account on the main YouTube platform.

Manage recommendations and search results

Users can also manage the recommendations they get by choosing to <u>view, delete, or</u> <u>turn on or off their YouTube watch and search history</u> whenever they want. Specifically,

if users do not want to see recommendations at all, they can turn off and clear their YouTube watch history. For users with YouTube watch history off and no significant prior watch history, the YouTube homepage will show the search bar and the Guide menu, with no feed of recommended videos.

All users, including those with supervised accounts on the main YouTube platform, can also tell YouTube when it is recommending something a user is not interested in. For example, buttons on the homepage and in the 'Up next' section allow users to filter and choose recommendations by specific topics. Users can also click on 'Not interested' and/or 'Don't recommend channel' to tell YouTube that a video or channel is not what a user wanted to see at that time, and YouTube will take that into consideration when generating recommendations for that viewer in the future. Additional information about how a user can manage their recommendation settings are outlined here in YouTube's Help Centre.

YouTube also introduced <u>safeguards reducing repeated recommendations</u> of certain content to specifically support the well-being and mental health of young people. For example, while some types of content may seem innocuous in a single view, experts have found that repeatedly viewing some types of content can be problematic for young people, such as content that displays social aggression, compares physical features or idealises some body types, fitness levels or weight over others.

Ad safeguards for children on YouTube

For users below 18 years of age, remarketing and other personalised targeting are specifically prohibited on YouTube Kids and YouTube's Supervised Experiences. We also have <u>safeguards</u> in place to prevent age-sensitive ad categories from being shown to children.

Blocking, reporting and turning off ads on YouTube

Users can easily customise the ads they see on YouTube and control the info used to personalise them via My Ad Center. Additionally, if a user thinks an ad they have been served is inappropriate or violative of Google's ad policies, the user can report it. More information can be found here.

Users also have the ability to turn off ads on YouTube by signing up for YouTube Premium, which allows users to watch videos without interruptions by ads before and during a video, including video overlay ads. Users who are YouTube Premium members will also not see third-party banner ads and search ads. Ad-free videos are supported across all devices and platforms where you can sign in with your Google Account—including on compatible smart TVs/gaming consoles and the YouTube and YouTube Kids mobile apps, if they're available in your location.

Paragraph 19: Children or their parents/ guardians must have access to tools that enable them to manage children's safety, and effectively minimise children's exposure to, and mitigate the impact of, harmful and/ or inappropriate content and

Please provide information on the measures, and include screenshots or evidence where possible.

Children:

Paragraph 19(a): Tools to effectively manage the content that children see and/or their experiences.

See our response to Paragraph 9 on YouTube Kids and Supervised Experiences.

unwanted interactions on the Service. These tools may include the following:

(a): Tools to effectively manage the content that children see and/or their experiences.

(b): Tools to:

- Limit the public visibility of children's accounts, including their profile and content;
- ii. Limit who can contact and/or interact with children's accounts; and
- iii. Limit location sharing

Paragraph 20: Unless the Service restricts access by children, children must be provided differentiated accounts whereby the settings for the tools to minimise exposure and mitigate impact of harmful and/or inappropriate content and unwanted interactions are robust and set to more restrictive levels that are age appropriate by default. Children or their parent/ guardians must be provided clear warnings of implications if they opt out of the default settings.

Children:

Paragraph 19(b): Tools to:

- . Limit the public visibility of children's accounts, including their profile and content:
- ii. Limit who can contact and/or interact with children's accounts; and
- iii. Limit location sharing.

See our response to Paragraphs 12(a) to (c), which also applies to users who are children. In addition, see also our response to Paragraph 9 above and Paragraph 20 below on the default privacy settings for users below the age of 18.

Please provide information on the measures, and include screenshots or evidence where possible.

Children:

YouTube Kids and Supervised Experience

For children under the age of 13, we provide differentiated account types through YouTube Kids and Supervised Experiences. See our response to Paragraph 9 on YouTube Kids and Supervised Experiences.

To enable a child to access Supervised Experiences on the main YouTube platform, parents must create a <u>supervised Google account</u> for them.

To enable a child to access YouTube Kids, parents have the option to set up the app in a signed out state, or to sign into the app and create a profile for their child (associated with the parent's Google account). Finally, if their child has a supervised Google account, parents can also set up YouTube Kids for their child's account.

In these account types, we provide more restrictive settings which include:

- Age restriction for certain types of content (see our response to Paragraph 9 and 17)
- Parental controls in both the <u>YouTube Kids experience</u> and the <u>Supervised</u>
 <u>Experiences</u>, where parents can block content
- YouTube Kids offers content settings that align to different ages. "Preschool" aligns with ages 4 and under. "Younger" aligns with ages 5–8. "Older" aligns with ages 9–12. Parents can also define profiles for their kids' viewing to handpick the content they have access to.
- Supervised Experiences offer content settings that align to different ages.
 "Explore" generally aligns with content ratings for viewers aged 9+, "Explore More" generally aligns with content ratings for viewers aged 13+, and "Most of YouTube" is where almost all videos on YouTube would be accessible except for

content marked as 18+ or otherwise not appropriate for viewers using supervised experience. Depending on the content setting the parent picks, their child can also view <u>comments on videos</u>.

Users are able to find out about the differences between YouTube Kids and Supervised Experiences, and therefore the implications of switching between the two types of experiences, <u>here</u>.

Age Restriction and Made for Kids Principles

In addition to the safeguards in place via YouTube Kids and Supervised Experiences, YouTube has the following safeguards in place to minimise exposure to age-inappropriate content:

- Age restriction: Please see our response in Paragraph 9 above for details.
- Made for Kids Principles: Please see our response in Paragraph 9 above for details.

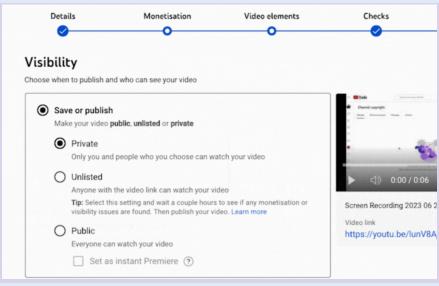
Safeguards limiting repeated recommendations of certain content

For users under the age of 18, YouTube introduced <u>safeguards limiting repeated</u> <u>recommendations</u> of certain content to specifically support the well-being and mental health of young people (see also our response to Paragraph 18 on managing content recommendations).

Privacy settings on video uploads

In order to help minimise unwanted interactions, YouTube also provides restrictive privacy settings. The default video privacy setting for videos uploaded by users aged 13–17 is private. For users under the age of consent, YouTube does not yet offer the ability to create a channel, and therefore those users cannot create videos or receive comments. YouTube also does not allow children under the age of 13 to appear in a live stream unless they are visibly accompanied by an adult. Channels not in compliance with this policy may temporarily or permanently lose their ability to live stream.

Implications of each privacy setting is clearly set out in the user interface, as shown in the screenshot below:



Section B: User Reporting and Resolution

Paragraph 23: Any individual must be able to report concerning content or unwanted interactions to the Service in relation to the categories of harmful and/or inappropriate content in paragraphs 4 and 17. In this regard, the reporting and resolution mechanism provided to end-users must be effective, transparent, easy to access, and easy to use.

Please provide information on the measures, and include screenshots or evidence where possible.

All End-users in Singapore:

Paragraph 23(a): End-users' reports must be assessed, and appropriate action(s) must be taken by the Service in a timely and diligent manner that is proportionate to the severity or imminence of the potential harm. In particular, timelines must be expedited for content and activity related to terrorism. Appropriate action(s) may include:

- Swiftly removing the reported content or restricting access to the reported content; and
- ii. Warning, suspending, or banning the account(s) that generated, uploaded, or shared the reported content.

Users can report content believed to violate YouTube's <u>Community Guidelines</u> using YouTube's <u>flagging feature</u>, which is available to signed-in users in Singapore via computer (desktop or laptop), mobile devices, and other surfaces. Details on how to report different types of content using YouTube's flagging feature is outlined in YouTube's <u>Help Centre</u>. Users can also report a <u>privacy complaint</u> if they believe someone has used their personal info without permission or submit a <u>legal complaint</u> if content identified may violate local law.

Please provide information on the measures, and include screenshots or evidence where possible.

All End-users in Singapore:

Paragraph 23(b): Where the Service receives a report that is not frivolous or vexatious:

- i. The end-user who submitted the report must be informed of the Service's decision and action taken with respect to that report without undue delay.
- ii. Should the Service decide to take action against the report content or account(s), the end user holding the account(s) that generated, uploaded, or shared the reported content must be informed of the Service's decision and action without undue delay.

Users in Paragraph 23(b)(i)

Users can <u>view their Reporting History</u> to check the status of videos they have reported on YouTube, which include:

- Live: Videos that are either not yet reviewed or that we decided don't violate YouTube Community Guidelines.
- Removed: Videos that have been removed from YouTube.
- Restricted: Videos that have been placed in a restricted state, like <u>age-restriction</u> or <u>Limited Features</u>.

Users in Paragraph 23(b)(ii)

When a creator's content is found to have violated Community Guidelines for the first time, they will typically get a warning with no penalty to their channel. They will have the chance to take a policy training to allow the warning to expire after 90 days. Creators will also get the chance to receive a warning in another policy category. If the same policy is violated within that 90 day window, the creator's channel will be given a strike. If a creator receives three strikes in the same 90-day period, their channel may be permanently removed from YouTube. In some cases, YouTube may terminate a channel for a single case of severe abuse, as explained in the Help Centre. YouTube may also remove content for reasons other than Community Guidelines violations, such as a first-party privacy complaint or a court order. In these cases, creators will not be issued a strike.

If a creator's channel gets a strike, they will receive an email, and can have notifications sent to them through their mobile and desktop notifications. The emails and notifications received by the creator explain the action taken on their content and which of YouTube's policies the content violated. More detailed guidelines of YouTube's processes and policies on strikes here.

Please provide information on the measures, and include screenshots or evidence where possible.

All End-users in Singapore:

Paragraph 23(c): The end-users referred to in sub-paragraphs (b)(i) and (b)(ii) must be allowed to submit requests to the Service for a review of the decision and action taken.

Users in Paragraph 23(b)(i)

YouTube provides all users the ability to report content for violation of YouTube's Community Guidelines ("CG flags"), and for violation of country-specific laws ("Legal Complaints"). If a user is dissatisfied with the outcome of their CG flag (i.e., when the video flagged has been assessed not to be policy-violative and remains available to view globally on YouTube), they can submit a request to YouTube to review the video in the form of a Legal Complaint here where the user believes the video violates Singapore law. YouTube will review the user's Legal Complaint and users will be informed of the outcome of their Legal Complaint via email. Where a user submits a Legal Complaint and the video is not initially actioned on legal grounds, the user may further request an internal appeal of YouTube's decision via the user's correspondence with YouTube.

Users in Paragraph 23(b)(ii)

YouTube gives creators the option to appeal video removals or restrictions in case they think we made a mistake. If a creator chooses to submit an appeal, it goes to human review, and the decision is either upheld or reversed.

The below appeals processes are available to all creators, which are outlined in the YouTube Help Centre:

- Appeal a Community Guidelines strike
- Appeal a Community Guidelines video removal
- Appeal the age restriction of a video
- Appeal playlist or thumbnail removals
- Appeal a channel termination

Section C: Accountability - Mandatory Information and Metrics		
Paragraph 24: End-users must	Please see our response to Paragraph 13 in Section A of this report.	
have access to clear and easily		
comprehensible information that		
enable them to assess the level of		
safety and related safety measures		
afforded by the Service and make		
informed choices.		
Paragraph 25(a): What steps the	Please see our responses in Sections A and B of this report.	
Service has taken to mitigate		
Singapore end-users' exposure		
to harmful or inappropriate		
content, including descriptions of		
specific measures		
that the Service has in place to		
enhance online safety for		
end-users in		
Singapore in relation to obligations		
in Sections A and B;		
Paragraph 25(b): How much and	Please see our response to Paragraph 26 in Section C of this report.	
what types of harmful or		
inappropriate content end-users in		
Singapore encounter on the		
Service		
Paragraph 25(c): What action(s)	Please see our response to Paragraph 26 in Section C of this report.	
the Service has taken on end-user		
reports.		

Section C: Accountability - Additional Information and metrics

Note: All data reported in this Section C is for the period of July 2023 to March 2024 ("**reporting period**"), unless otherwise stated.

Paragraph 26(a):

The number and types of end-user reports received from end-users in Singapore and the number and types of harmful and inappropriate content removed as a result of end-user reports.

(1) Number of videos flagged from a Singapore IP address for potential Community Guidelines ("CG") violations

During the reporting period, **396,001** video flags were reported by end-users with a Singapore IP address for potentially violating YouTube's CGs in respect of the following reasons:

Table 1.1 End-user flags from Singapore IP address, by reason

Reason	Number of flags received
Spam or Misleading	197,009
Sexual	80,389
Violent or Repulsive	38,103
Promotes Terrorism	32,971
Harmful or Dangerous Acts	28,449
Child Abuse	19,080

Note:

- The above categories of reasons for video flags are based on YouTube's internal categorisation, which does not exactly mirror, but broadly covers, the harmful content categories set out in Paragraph 4 of the Code of Practice.
- (2) Number of videos uploaded from a Singapore IP address that were removed for CG violations, broken down by YouTube's CG categories

During the reporting period, YouTube removed **1,974,552** videos that were uploaded from a Singapore IP address for violating YouTube's CGs in respect of the following reasons:

Table 1.2 Videos removed that were uploaded from a Singapore IP address, by reason

Reason	Number of videos removed
Harmful or Dangerous	1,956,180
Child Safety	6,917
Nudity or sexual	3,731
Harassment and Cyberbullying	3,613

Promotion of violence and violent extremism 657 Misinformation 441	Violent or Graphic	3,013
Misinformation 441	Promotion of violence and violent extremism	657
	Misinformation	441

Note:

- The above categories of reasons for video removal are based on YouTube's internal categorisation, which does not exactly mirror, but broadly covers, the harmful content categories set out in Paragraph 4 of the Code of Practice.
- (3) Number of videos uploaded from a Singapore IP address that were removed for CG violations, where the source of first detection was a user flag

During the reporting period, YouTube removed **9,476** videos that were uploaded from a Singapore IP address for violating YouTube's CGs, where the source of first detection was an end-user flag.

Note:

- The above data is based on video removals across all CG categories.
- (4) Number of legal complaints received from Singapore users related to the Singapore Code of Practice for Online Safety, and number of videos removed or geoblocked as a result of such legal complaints

Table 1.3 Legal complaints received from Singapore users related to the Singapore Code of Practice for Online Safety during the reporting period

Number of complaints received		Median time taken to take action (if any) on complaints
6	N/A	N/A

Notes:

- "Number of complaints received" are based on complaints that YouTube received during the reporting period via our "Other Legal Complaint" webform, citing the "Singapore Code of Practice for Online Safety".
- "Number of items removed" is based on items that YouTube, after reviewing the complaint, removed as a result of such a complaint.
- "Number of items removed" and "Median time taken to take action on complaints" are reflected as "N/A" above, because upon YouTube's review of the complaints, the basis of each complaint was either assessed to be unrelated to the Code of Practice or the content was already removed from YouTube for separate reasons.

Paragraph 26(b): The time between the Service receiving end-user reports from end-users in Singapore on harmful and inappropriate content and taking action (if any) as an aggregate. See our response to Paragraph 26(a) above. Paragraph 26(c): The number and types of harmful or inappropriate content proactively removed by the Service that are:

- i. Accessible by end-users in Singapore; and
- (1) Global number of videos removed for CG violations, where the source of first detection was automated flagging

During the reporting period, YouTube removed **24,405,862** videos where the source of first detection was automated flagging.

- ii. Originated from Singapore
- (2) Number of videos uploaded from a Singapore IP address that were removed for CG violations, where the source of first detection was automated flagging

During the reporting period, YouTube removed **1,978,602** videos uploaded from a Singapore IP address where the source of first detection was automated flagging.

Note:

• The above data is based on video removals across all CG categories.

Paragraph 26(d):

The number of accounts suspended or banned in Singapore, and the reasons for suspending or banning accounts in relation to the categories of harmful and inappropriate content.

(1) Global number of channels removed for CG violations

During the reporting period, YouTube removed **2,800,719** channels for violating YouTube's CGs in respect of the following reasons:

Table 2.1 Global number of channels removed, by reason

Reason	Number of channels removed
Nudity or Sexual	1,497,488
Misinformation	504,795
Child Safety	343,500
Harmful or Dangerous	192,539
Promotion of Violence and Violent Extremism	112,030
Harassment and Cyberbullying	96,672
Violent or Graphic	53,695

- The above categories of reasons for channel removal are based on YouTube's internal categorisation, which does not exactly mirror, but broadly covers, the harmful content categories set out in Paragraph 4 of the Code of Practice.
- (2) Number of Singapore channels removed for CG violations, broken down by YouTube's CG categories

From October 2023 to March 2024, YouTube removed **20,739** Singapore channels for violating YouTube's CGs in respect of the following reasons:

Table 2.2 Number of Singapore ch	nannels removed, by reason
----------------------------------	----------------------------

Reason	Number of channels removed
Harmful or Dangerous	13,853
Nudity or Sexual	5,982
Misinformation	609
Child Safety	182
Harassment and Cyberbullying	55
Promotion of Violence and Violent Extremism	33
Violent or Graphic	25

Note:

- The above categories of reasons for channel removal are based on YouTube's internal categorisation, which does not exactly mirror, but broadly covers, the harmful content categories set out in Paragraph 4 of the Code of Practice.
- The above data is in respect of the period of October 2023 to March 2024, as this data was only available from October 2023 onward.