

# TikTok response to Ofcom consultation on draft video-sharing platform guidance

## 1. Executive Summary

TikTok is grateful for the opportunity to provide feedback on the latest video-sharing platform guidance published by Ofcom, “Guidance for providers on measures to protect users from harmful material”. We recently notified Ofcom that the TikTok service to UK users falls within Ofcom’s jurisdiction over video sharing platforms (“VSPs”) and have warmly welcomed the constructive engagement with the Ofcom team as it has developed guidance in this area.

TikTok’s top priority is to promote a safe and positive experience so that all our users can feel free and empowered to express their creativity.

In September 2020 TikTok responded to Ofcom’s call for evidence on video-sharing platform regulation, in which we outlined in detail our approach to ensuring trust and safety on our platform. This submission covered a wide range of topics, including our Terms of Service (<https://www.tiktok.com/legal/terms-of-service>) and Community Guidelines (<https://www.tiktok.com/community-guidelines>), our approach to content moderation and reporting, our age-assurance policies, measures to protect young people, and our commitment to transparency.

We are constantly looking for new ways to make our platform safer for our users. Since September 2020 we have introduced a number of significant new measures as part of our ongoing commitment to protecting and empowering our users. These changes are described in this response and include:

- **October 2020:** we implemented improved notifications to provide clarity to users around content removals
- **December 2020:** we strengthened our Community Guidelines to better support the well-being of our community
- **January 2021:** we changed the accounts of users under the age of 16 to private by default
- **February 2021:** we introduced new prompts to encourage users to consider the impact of their content before they share it and we published our Transparency Report for H2 2020
- **March 2021:** we launched our European Safety Advisory Council and introduced new features to control comments and prompts to promote kindness
- **April 2021:** we announced our European Transparency and Accountability Centre
- **May 2021:** we launched a new media literacy campaign to promote critical thinking

At TikTok we support the goals of Part 4B of the Communications Act 2003 (“the Act”) which implements the revised AVMS Directive (“the Directive”). We support Ofcom’s stated approach that the guidance should give VSPs flexibility in how they protect their users, bearing in mind the diversity of the sector and technological innovation. We believe that an overly prescriptive approach would not achieve Ofcom’s stated intentions.

We believe it is important to acknowledge that different measures may be appropriate for different types of content, and even different VSPs given their unique nature. For example, hash lists can be integrated and other technologies can be built to proactively detect and remove certain illegal content such as CSAM. On the other hand, cyberbullying is often a pattern of conduct which individual posts on their own may not reveal, and technology-based tools may not be as effective at detecting such content.

To this theme, we agree with the importance of measuring the effectiveness of actions that VSPs implement. We believe the effectiveness of measures is also best determined by individual VSPs who will have the clearest picture of their users and challenges at hand (as well as technical capabilities), subject to high-level principles set out by Ofcom.

TikTok’s submission is focused on two main areas: the first outlines TikTok’s evolving approach to safety, including the measures implemented since the last consultation on VSP guidance in September 2020. The second provides views on the questions set out in the latest consultation document. In this second section we would like to particularly highlight our calls for:

- Avoiding an overly prescriptive approach to what may constitute “restricted material” and in particular “material that might impair the physical, mental or moral development of persons under the age of 18”, taking into account that the Act does not give examples of restricted material; moderators have limited scope to understand the context and background of some content; and there are risks of over-moderation created by expecting moderators to adjudicate on user intent.
- Recognition of the complexity and subjectivity of designing and implementing policies relating to restricted material that seek to balance protective measures with broader rights to acquire knowledge, connect with others, seek enjoyment and self-expression.
- Acknowledgment that there are several different approaches to content rating that may be available to VSPs, including platform-designed classification, classification by standards bodies, and user-generated classification and confirmation that each VSP is best placed to determine which particular method is likely to be most effective on its platform.
- Confirmation that users’ privacy rights may place limits on how certain measures are implemented, for example in relation to media literacy measures, there are potential restrictions on what quantitative data can be collected or shared on users, tools and campaigns.

- Reassurance that as the “Additional steps to protect users” (which we agree may be very helpful) are not required by the Act, they are illustrative only and that VSPs will not be required to implement these in order to comply with the Act.
- In relation to the dispute resolution procedure, reconsideration of the approach, in particular in relation to the types of dispute that are eligible to be considered under the procedure and the procedure that may be followed, to ensure that it is proportionate and consistent with the legislation.

## 2. TikTok’s evolving approach to safety

At TikTok, our mission is to inspire creativity and bring joy. TikTok’s top priority is to promote a safe and positive experience so that all our users can feel free and empowered to express their creativity.

There are three main ways we do this:

1. Our Terms of Service (<https://www.tiktok.com/legal/terms-of-service>) and Community Guidelines (<https://www.tiktok.com/community-guidelines>) reflect our values and establish the kind of behaviour we expect from our community. We enforce these rules using a combination of cutting-edge technology and thousands of safety experts based around the world.
2. We develop robust safety policies and features, including restricting direct messaging to users who are 16 and older and allowing parents to pair their account with their teen’s account so that they can enable a variety of content and privacy settings through Family Pairing.
3. We collaborate with industry partners to make the digital world safer for everyone. We work with safety organisations such as Internet Matters and are signatories to industry-wide initiatives such as the Voluntary Principles to Counter Online Child Sexual Exploitation and Abuse, the Code of Practice on Disinformation and the Code of Conduct on Countering Illegal Hate Speech Online.

At TikTok, we are constantly looking for new ways to make our platform safer for our users. Since our previous submission, we’ve introduced a number of significant new measures as part of our ongoing commitment to protecting and empowering our users. These are outlined below, and if helpful we would be delighted to brief Ofcom on these in greater detail.

### **Adding clarity to content removal**

In October 2020 we [rolled out improved notifications](#) to users around content removal. By being more transparent with our content enforcement and educating users we aim to ensure that our



Community Guidelines are enforced uniformly and evenly, reducing misunderstandings about content on our platform.

Now, when content is removed for violating our policies, we let the creator know which policy was violated leading to the removal, and provide the creator with the ability to appeal the decision. We believe this helps foster greater understanding of the kind of positive content and welcoming behavior that makes our community thrive.

Explaining our enforcement actions while reminding people of our policies has helped reduce the rate of repeat violations, and visits to our Community Guidelines have nearly tripled. We've also seen a 14 percent reduction in requests from users to appeal a video's removal.

### **Updated our community guidelines**

In December 2020 we [announced updated Community Guidelines](#) with additional details about what's allowed on TikTok. The update adds more specific examples to the categories of content the guidelines cover, based on behaviour we've seen on the platform, feedback we've heard from our community, and input from academics, civil society organizations, and our Content Advisory Council.

While much of this content was already covered by our previous guidelines, below are some of the key areas we've strengthened to better support the well-being of our community:

- **Self-harm and eating disorders:** Our updated guidelines incorporate feedback and language used by mental health experts to improve our policies on self-harm and suicide content and avoid normalizing self-injury behaviors. Our policy on eating disorder content has additional considerations to prohibit normalizing or glorifying dangerous weight loss behaviors.
- **Bullying and harassment:** We recognise the burden victims of abuse often face in managing their online presence. We've bolstered our policies on bullying and harassment and our guidelines are now more explicit about the types of content and behaviors that aren't welcome on TikTok, including doxxing, cyberstalking, and a more extensive policy against sexual harassment.
- **Dangerous and harmful activities:** The safety of everyone in our community is of utmost importance, especially the well-being of youth. In line with our existing dangerous acts policy, we work to either limit, label, or remove content that depicts dangerous acts or challenges. Now, we've added a harmful activities section to our minor safety policy to reiterate that content promoting dangerous dares, games, and other acts that may jeopardize the safety of youth is not allowed on TikTok.
- **Violence and extremism:** TikTok stands firmly against violence, both online and off. We've updated our previous dangerous individuals and organizations policy to focus more holistically on the issue of violent extremism. Our guidelines now describe in



greater detail what's considered a threat or incitement to violence and the content and behavior we prohibit.

### **Strengthening privacy and safety for youth on TikTok**

In January 2021 we [announced changes](#) for users under the age of 18 aimed at driving higher default standards for user privacy and safety. In particular, we changed the default privacy settings for accounts registered with an age of 13-15 to private, meaning only someone who the user approves as a follower can view their videos.

Additional changes we rolled out to promote a safe experience for our younger users include:

- Tightening the options for commenting on videos created by those under age 16. These users can now choose between Friends or No One for their account; the Everyone comment setting has been removed.
- Changing Duet and Stitch settings so no one can make duets or stitches (co-creation features) with content created by people under 16. For users aged 16-17, the default setting for whether duets and stitches can be made with their content is set to Friends.
- Only allowing downloads of videos that have been created by users 16 and over. Other users can decide whether they want to allow downloads of their videos, though for users aged 16-17 the default setting will now be changed to Off.
- Setting "Suggest your account to others" to Off by default for users under 16.

These changes built on previous updates we've made to promote child safety, including:

- Restricting direct messaging and hosting live streams to accounts 16 and over.
- Restricting the buying, sending, and receiving of virtual gifts to users 18 and over.
- Enabling parents and guardians to work with their teen to set guardrails on their teen's TikTok experience through our Family Pairing features.

### **New prompts to help people consider before they share**

As part of our efforts to limit the spread of potentially misleading information TikTok [introduced its Know the Facts feature to encourage people to pause before sharing videos containing unverified information](#) in February 2021.

We remove misinformation as we identify it and partner with fact checker Logically in the UK to help us assess the accuracy of claims made in certain kinds of content. If fact checks confirm the claims in the content to be false, we'll remove the video from our platform. However, sometimes fact checks are inconclusive or it is not possible to verify whether the content of videos is truthful or not, especially during unfolding events. In these cases, a video may become



ineligible for recommendation into users' For You feeds to limit the spread of potentially misleading information.

Now, if a user views a piece of content that has been received by fact checkers but cannot be conclusively validated, they will see a banner, and the video's creator will also be notified that their video was flagged as unsubstantiated content. If a viewer passes the banner and attempts to share the flagged video, they'll see a prompt reminding them that the video has been flagged as unverified content. This additional step requires a pause for people to consider their next move before they choose to "cancel" or "share anyway."

We've designed this feature to help our users be mindful about what they share and to dissuade users from sharing content that may not be accurate. When we tested this approach we saw that there was a reduction in the rate at which viewers shared videos by 24%, while likes on such unsubstantiated content also decreased by 7%.

### **TikTok's European Safety Advisory Council**

In March 2021 we [launched our Safety Advisory Council for Europe](#) to support us in our efforts to make TikTok as safe as it can be.

The Council brings together leaders from academia and civil society from all around Europe. Each member brings a different, fresh perspective on the challenges we face and members will provide subject matter expertise as they advise on our content moderation policies and practices. Not only will they support us in developing policies that address the challenges we face today, they will also be forward-looking and help us to identify emerging issues that affect TikTok and our community in the future.

We are humbled that these industry leaders in their respective fields are joining the Council and as the challenges that we face evolve. In addition to the nine inaugural Council members TikTok will add additional members from more countries and different areas of expertise in the future as we continue our journey to help make TikTok a place where joy and creativity can thrive.

### **TikTok's European Transparency and Accountability Centre**

We believe that being open about what's happening behind the scenes is crucial to helping people better understand TikTok and our work, as well as maintaining the trust of our community, policymakers and the broader public.

In April 2021 we announced our new [European Transparency and Accountability Centre](#). Once operational, the Centre will provide experts with an opportunity to visit and see first-hand how teams at TikTok go about the critically important work of securing our community's safety, data, and privacy.

Through this direct observation of our practices, experts will have an opportunity to learn about our moderation systems, processes, and policies. This includes exploring:



- How we use technology to keep our community safe
- How our trained content review teams make decisions about content based on the policies in our Community Guidelines
- The way human reviewers supplement our moderation efforts using technology to help catch potential violations of our policies
- Ultimately, how the content allowed on the platform aligns with our values

In addition to content moderation, the Centre will offer detailed insight into our recommendation technology, and our commitment to data privacy and security.

The landscape we operate in is rapidly evolving and it's our hope that visitors will be able to learn more about our work, but importantly, also provide candid feedback about what they see and hear. No system, policy, or practice is flawless, and we are committed to constant improvement.

### **New tools to promote kindness on TikTok**

At TikTok, our goal is to promote a positive environment where people support and lift each other up. To help foster kindness in our community, in March 2021 we [rolled out two new features](#): the first feature gives creators more control over the comments on their videos, and the second prompts people to reconsider posting unkind or inappropriate comments.

With our new Filter All Comments feature, creators can decide which comments will appear on their videos. When enabled, comments aren't displayed unless the video's creator approves them using the new comment management tool. This feature builds on our [existing collection of comment controls](#) that allow people to filter spam and offensive comments, and specific keywords. We know that comments are an important way for community members to connect with creators, and we'll continue to develop ways to promote respectful discussion.

In addition to empowering creators with more tools, we also want to encourage our community to treat everyone with kindness and respect. A new comment prompt now asks people to reconsider posting a comment that may be inappropriate or unkind. It also reminds users about our [Community Guidelines](#) and allows them to edit their comments before sharing.

### **#FactCheckYourNewsFeed Media Literacy Campaign**

At TikTok, we take the responsibility to help educate our users seriously, which means making sure that everyone has access to good, accurate information whenever they need it. To achieve this, it's important that our users are able to correctly identify what they are watching and to know if they are being given potentially inaccurate or harmful information.



In May 2021 we [collaborated with creators, public figures and digital media literacy charity The Student View](#) to roll out a series of videos on the platform aimed at keeping people on our platform safe and helping them to think critically.

The campaign has begun with a series of videos on critical thinking and news literacy. The first phase of videos will educate on how to analyse and evaluate information, build an argument, recognise fake news and misinformation, and consume a balanced 'news diet'.

As the campaign progresses, it will also educate on vaccines misinformation, financial literacy, consuming diet/exercise advice and challenging Us versus Them narratives.

We have a microsite on the Discover tab of the app that will house the content and key messages for the campaign. As we launch the topics every few weeks, new tabs will be added.

We are also making sure that users who search for these topics get these videos presented to them via a search intervention that links to a banner and the microsite. All users in the UK were also sent a push notification encouraging them to view the critical thinking content.

### **Guardian's Guide**

At TikTok, we recognise the challenges that parents and guardians face parenting in the digital age with countless devices and apps available to families today. Keeping teens educated and safe requires collaboration among parents, experts, governments, and platforms.

In that spirit, we have [rolled out a refreshed Safety Centre](#) with new guides and resources aimed at supporting digital safety and security conversations among families. Our goal is to make it easy for parents to understand how TikTok works and the tools they can use to create the best experience for their family.

This includes the Guardian's Guide, a one-stop-shop to learn all about TikTok including how to get started on the platform, our safety and privacy tools, such as our Family Pairing features, and additional resources to address common online safety questions.

We understand that every teen and family is different, which is why we offer a range of choices and settings to empower parents and caregivers who are looking for easy ways to engage in their teen's life online.

### **TikTok's H2 2020 Transparency Report**

Last but not least, in February we published our [fourth and latest Transparency Report](#). This Transparency Report covers the second half of 2020 (1 July - 31 December).

The Transparency Report provides visibility into the volume and nature of content removed for violating our Community Guidelines or Terms of Service, with additional insight into our work to counter misinformation related to COVID-19 and elections. It also includes how we respond to





law enforcement requests for information, government requests for content removals, and intellectual property removal requests.

We've added a number of new data points to this report in our effort to be ever more transparent, including:

- Accounts removed
- Spam accounts and videos removed
- Videos restored after they were appealed by the video's creator
- Specific policy insights
- Ads rejected for violating our advertising policies

Below are some key insights from the report:

- 89,132,938 videos were removed globally in the second half of 2020 for violating our Community Guidelines or Terms of Service, which is less than 1% of all videos uploaded on TikTok.
- 92.4% of these videos were removed before a user reported them, 83.3% were removed before they received any views, and 93.5% were removed within 24 hours of being posted.
- 6,144,040 accounts were removed for violating our Community Guidelines.
- 9,499,881 spam accounts were removed along with 5,225,800 spam videos posted by those accounts. We prevented 173,246,894 accounts from being created through automated means.
- 3,501,477 ads were rejected for violating advertising policies and guidelines.

### **3. TikTok's response to consultation questions**

#### **Definitions**

- 1. Do you have any comments on Section 3 of the draft guidance on harmful material and related definitions?*

TikTok's Terms of Service and Community Guidelines prohibit users from sharing the types of content outlined in Section 3 of the draft guidance and in many places go further than the categories outlined.

#### **Relevant harmful material**

We note that the definition of "relevant harmful material" includes material likely to incite violence or hatred and material the inclusion of which would be a criminal offence under laws



relating to terrorism, child sexual abuse material, racism and xenophobia. Each of these types of content are prohibited by TikTok's Terms of Service and Community Guidelines.

- TikTok's Terms of Service state that users may not:
  - "intimidate or harass another, or promote sexually explicit material, violence or discrimination based on race, sex, religion, nationality, disability, sexual orientation or age"
  - "use the Services to either intentionally, recklessly or negligently upload, transmit, distribute, store or otherwise make available... any material which does or may infringe applicable laws or which infringes someone else's rights"
  
- TikTok's Community Guidelines also forbid:
  - Hateful behaviour, including attacks on the basis of protected attributes, slurs, and hateful ideologies
  - Violent extremism including but not limited to threats and incitement to violence and dangerous organisations and individuals who promote or are engaged in violence
  - Sexual exploitation of minors, nudity and sexual activity involving minors, harmful activities by minors, physical and psychological harm of minors, grooming behaviour and crimes against children

We welcome and support the obligation placed on VSPs to take appropriate measures to protect all users from these categories of content. TikTok is committed to keeping its platform safe and agrees that one of the aims of effective implementation should be to prevent users from encountering harmful material. However, given the volume of content that is uploaded to TikTok and that no VSP's content moderation is infallible, we welcome Ofcom's acknowledgment that the implementation of measures will not always lead to a total eradication of harmful material from a platform.

We also strongly agree that each VSP is best placed to determine what measures will be appropriate for it to implement in order to prevent users from encountering harmful material, or reduce the prevalence of it, or restrict access to it.

We note that the draft guidance recommends that VSPs "consider the context in which the potentially harmful material is presented, such as the intention of the inclusion of the material as well as any information it has about the user uploading the material", to avoid the removal of harmful material contained in news reports which is necessary for informing audiences. Although we support the intention of protections for journalistic content, we note that when moderating content at scale, it can be difficult or at times impossible to determine the intent behind a user's content. Further, sometimes there will be clear information to indicate that content is being shared from a reliable news source, but on other occasions that information will not be available. For this reason we agree that giving users the opportunity to appeal take-down decisions is important.

## Restricted material

We note that the definition of “restricted material” includes videos which have or would be likely to be given an R18 certificate; videos containing material determined not suitable for BBFC classification or material it is reasonable to expect would be determined as not suitable for classification and material that might impair the physical, mental or moral development of under 18s.

TikTok’s Terms of Service and Community Guidelines prohibit users from sharing categories of content that would clearly cause harm to under 18s.

- TikTok’s Terms of Service state that users must not share:
  - any material that would constitute, encourage or provide instructions for a criminal offence, dangerous activities or self-harm;
  - any material that is deliberately designed to provoke or antagonise people, especially trolling and bullying, or is intended to harass, harm, hurt, scare, distress, embarrass or upset people;
  - any material that contains a threat of any kind, including threats of physical violence.
  
- TikTok’s Community Guidelines state that we do not allow:
  - content that is gratuitously shocking, graphic, sadistic, or gruesome or that promotes, normalizes, or glorifies extreme violence or suffering
  - content depicting, promoting, normalizing, or glorifying activities that could lead to suicide, self-harm, or eating disorders. We also do not permit users to share content depicting them partaking in, or encouraging others to partake in, dangerous activities that may lead to serious injury or death
  - members of our community being shamed, bullied, or harassed
  - nudity, pornography, or sexually explicit content
  - activities that perpetuate the abuse, harm, endangerment, or exploitation of minors
  - activities that may undermine the integrity of our platform or the authenticity of our users

TikTok supports the legislative goal of protecting younger users from clearly harmful content. However, what might impair the “physical, mental or moral development” of a child will depend on various factors including the age, the developmental maturity and pre-existing vulnerabilities of the child, as well as cultural nuances and the context of the content. As noted in the draft guidance, the legislation does not specify any examples of restricted material.

We note that the draft guidance also encourages VSPs to balance protective measures with recognition of the benefit to young people that a service might offer such as acquiring

knowledge, connecting with others and seeking enjoyment and self-expression. Whilst we agree that a balance needs to be struck, we believe there must be recognition of the complexity and subjectivity in creating and implementing policies to apply to content that may impair the physical, mental or moral development of some under 18s, in comparison to clearly definable, illegal content (which Ofcom categorises as “harmful material”).

In that respect, we note that the draft guidance includes a number of potential harms that could be relevant to consider when determining which measures it may be appropriate to take to protect under 18s from material that might impair their physical, mental or moral development. These include “body image” and “manipulation intended to harm”. We appreciate that Ofcom’s draft guidance suggests only that VSPs may find it relevant to *consider* these (and other harms identified in a research study) but would welcome confirmation that Ofcom will not hold VSPs to these standards, which are not prescribed by the legislation.

We note that in the final part of Ofcom’s draft guidance which considers the impact of the VSP regime, Ofcom recognises that there is the potential for VSPs subject to UK jurisdiction to lose market share to VSPs in other jurisdictions if regulatory standards elsewhere are lower but that this potential impact is mitigated because the UK adhered to the minimum requirements of the Directive. We would welcome Ofcom’s confirmation that it will maintain the approach of adhering to the minimum requirements of the Directive. We note that the Online Safety Bill will introduce broader obligations in relation to the protection of minors and consider this would be the appropriate time to introduce the higher standards rather than requiring VSPs under Ofcom’s jurisdiction to comply with higher standards ahead of those wider obligations becoming law.

Furthermore, in considering the examples of harms that might impair the physical, mental or moral development of under 18s, it is important to recognise that moderators often have limited scope to determine the intention of the person sharing the material and do not always have knowledge of the relevant context in which a user shares content. It is not reasonable to expect moderators to act as adjudicators as to a user’s intent when they do not have access to critical context and background. For example, it may not be possible for a moderator to make a judgement as to whether a piece of content has been shared with the intention of triggering body image related harm to a teenager.

The imposition of legal obligations in relation to ill-defined harms, or more complex harms that require additional context and background not available to VSPs raises the risk of over-moderation of legitimate content, potentially impacting users’ right to freedom of expression and their right to receive information.

For these reasons we consider it is important that the guidance recognises that this category of content is subject to interpretation, does not apply uniformly to all under 18s and may change over time according to societal developments. VSPs are best placed to develop policies that set the rules for their platform to ensure under 18s are kept safe. TikTok has experts in minor safety who provide this crucial insight into policy development and review.

## Terms and Conditions

2. *Do you have any comments on the draft guidance about measures which relate to terms and conditions, including how they can be implemented?*
3. *Regarding terms and conditions which prohibit relevant harmful material, do you have any comments on Ofcom's view that effective protection of users is unlikely to be achieved without having this measure in place and it being implemented effectively?*
4. *Do you have any comments on Ofcom's view that, where providers have terms and conditions requiring uploaders to notify them if a video contains restricted material, additional steps will need to be taken in response to this notification to achieve effective protection of under-18s, such as applying a rating or restricting access?*

Our Community Guidelines prohibit content that we would consider to constitute relevant harmful or restricted material. All users are required to comply with our Community Guidelines under our Terms of Service. Other VSPs may have similar arrangements. We agree that this measure is fundamental to the VSP regime.

We agree that terms and conditions need to be implemented in such a way to meet the requirement of protecting users and that in practice this is likely to mean having robust processes in place to ensure they are enforced. As noted above, TikTok uses a combination of automated and human moderation to enforce our Terms of Service and Community Guidelines.

However, a requirement on VSPs to amend terms of service so that uploaders must declare their intention to upload restricted material may not have a material impact on some providers since such content may be prohibited in any event. We agree that this measure alone is unlikely to be an effective measure for many VSPs to protect users from harmful and restricted content.

As discussed in the following section, there is also a question of user-accuracy when it comes to assessing their own content, particularly given the potential subjective nature of some restricted material.

## Reporting, Flagging & Rating Mechanisms

5. *Do you have any comments on the draft guidance about reporting or flagging mechanisms, including on Ofcom's view that reports and flagging mechanisms are central to protecting users?*
6. *Do you have any comments on the draft guidance about systems for viewers to rate harmful material, or on other tagging or rating mechanisms?*

Our latest Transparency Report (<https://www.tiktok.com/safety/resources/transparency-report-2020-2>) confirms that 92.4% of removed videos were actioned before a user reported them, because they were identified by our automated and human moderation. However, TikTok strongly agrees that reporting mechanisms are crucial to protecting users, and this is particularly the case for certain categories of content that may not always be caught by automated moderation, or where the context behind certain content is not clear. For example, user reporting is an important tool by which users can make platforms aware of content that may not on the face of it appear to a third party to violate Community Guidelines but may constitute bullying because of a pattern of behaviour and/or background context. We encourage users to use TikTok's in-app and web-app tools to report any content to us that they believe violates our Community Guidelines and all reports are considered by our specialist moderation teams. The user who reports the content will be informed of the outcome of their report and if a creator's content is taken down, the creator will be informed and told the reason for the take down.

In relation to the rating of content, we note that the Act refers to "easy to use systems allowing viewers to rate harmful material" but that the draft guidance indicates that "this measure is primarily about allowing viewers on a VSP to apply ratings to restricted material." For the reasons set out below, we agree that it may not be useful to ask users to rate harmful material which may already be prohibited by the VSP. With regards to the rating of restricted material, we believe that there are several different approaches that may be available to VSPs, including platform-designed classification, classification by standards bodies, and user-generated classification (and the user may be the uploader or viewer).

We have some reservations about user-generated classification in isolation from other classification approaches. As the draft guidance itself acknowledges, user-driven rating of restricted content is not widespread and comes with a number of risks. Whether content constitutes restricted material (and in particular what the "age-appropriate rating" of a piece of content should be) is subjective and even well-intentioned users may have low accuracy in rating content. Some users rating content as a creator may attempt to "game" the system to benefit the reach of their own content or disadvantage others. Maintaining high-accuracy may be particularly difficult for users of TikTok (especially viewers), whose main experience of content rating will be from long-form content. There is potential for inconsistency and inaccuracy which would likely require substantial resources to correct.

We note that the draft guidance states that VSPs may rely on existing age ratings frameworks in order to rate content on their platform. Whilst existing frameworks are valuable, they may not always neatly map across to the short-form nature of video content on TikTok.

It would be helpful for the guidance to acknowledge that there are a number of ways that content may be rated or classified and that each VSP is best placed to determine which particular method is likely to be most effective on its platform.

## Age Assurance and Parental Controls

7. *Do you have any comments on the draft guidance about age assurance and age verification, including Ofcom's interpretation of the VSP Framework that VSPs containing pornographic material and material unsuitable for classification must have robust age verification in place?*
8. *Do you have any views on the practicalities or costs relating to the implementation of robust age verification systems to prevent under-18s from accessing pornographic material and material unsuitable for classification? Please provide evidence to support your answer wherever possible.*
9. *Do you have any comments on the draft guidance about parental control systems?*

TikTok supports the principle of a risk-based approach to age-assurance measures which includes ensuring services are used only by those who meet the minimum age requirement and providing an age appropriate experience on the service. Age assurance is an industry-wide challenge. We know there is more work to be done and we are committed to working collaboratively with industry peers, regulators, and key stakeholders to find solutions.

We hope that Ofcom also recognises this and that this informs Ofcom's approach to regulating VSPs' measures for protecting young people online. In addition, we ask that Ofcom ensure that its recommendations are aligned with other existing regulatory measures designed to protect young people. TikTok is actively working towards compliance with ICO's Age Appropriate Design Code which requires services to establish age with a level of certainty that is appropriate to the risks; and also other similar proposals, for example the Irish DPC's Fundamentals for a Child-Oriented Approach to Data Processing.

We welcome Ofcom's guidance that VSPs should consider how proportionate their age assurance measures are in preventing access to under 18s based on the harm that material might cause. We are supportive of Ofcom's acknowledgement of the value of neutral age gates as part of a meaningful age assurance strategy, and the role for trust-based tools like parental controls as potential alternatives to age assurance. We also welcome Ofcom's acknowledgement of the need to consider exclusion risks in the development of age assurance systems and the importance of avoiding disrupting the user experience.

As noted above, TikTok's Community Guidelines prohibit nudity, pornography, or sexually explicit content. However, TikTok agrees that this type of restricted material which has the most potential to harm the physical, mental or moral development of under 18s should be subject to the strictest access control measures. In relation to other categories of restricted material, TikTok welcomes Ofcom's recognition of the limits to how far VSPs are able to prevent age-inappropriate material from appearing on a service, and the difficulties with verifying the age of under-18s using a service.

With regards to parental controls, the draft guidance states that the success of parental control systems will “depend on the level of trust in the parent/carer-child relationship”. While this is likely to be true, it is not something that VSPs have control or influence over. Despite platforms’ best efforts it is this relationship that is most likely to impact the uptake of these controls. Therefore, it should not be taken into consideration when considering VSPs’ compliance with obligations under the regime.

TikTok is proud of its Family Pairing feature which enables a parent or guardian to pair their account with their teen’s to set appropriate guardrails. However, in designing this feature we are mindful of the need to take into account the rights of children. We note that Ofcom has set out principles that VSPs are encouraged to take into account in designing and implementing protection measures which include that the measure should be “fair”: “measures should be designed and implemented in a way that does not unduly discriminate between users, introduce bias or result in inconsistent application.” We would encourage Ofcom to reflect in its guidance the importance of VSPs respecting the rights of the child when providing parental controls.

## Complaints and Dispute Resolution

*10. Do you have any comments on the draft guidance about the measure regarding complaints processes or on the regulatory requirement to provide for an impartial dispute resolution procedure?*

### Complaints processes

The Act includes the following as one of the appropriate measures that VSPs may decide to implement:

*“In relation to the implementation of the measures set out in paragraphs 6 and 7, [VSPs may] establish and operate a complaints procedure which must be transparent, easy to use and effective, and must not affect the ability of a person to bring a claim in civil proceedings”* [emphasis added]. Paragraphs 6 and 7 relate to flagging and reporting mechanisms and the explanations associated with them; systems for users to rate harmful material; age assurance systems; and parental control systems. Therefore it is these measures to which the complaints procedure should apply.

However, the latest draft guidance states that VSPs should have a complaints process that allows users to *“raise issues with the platform”*, and that *“although the complaints process in the VSP Framework is limited to these measures [i.e. the narrower measures envisaged by the Act], we consider it best practice for providers to have a process that covers all aspects of user safety and strongly recommend that providers consider implementing such a process”* [emphasis added]. Whilst TikTok agrees with the importance of VSPs providing the opportunity for users to



raise concerns and issue complaints, we would welcome confirmation in Ofcom's final guidance that it will not hold VSPs to standards which are not required by the Act.

Ofcom's draft guidance also suggests that the "*likely timeframe of the complaint process*" could be provided to users wishing to make a complaint. While we consider this reasonable in principle, it is important that the final guidance acknowledges the fact that some complaints will be more complex and time-consuming to handle than others, making precise timeframes difficult for VSPs to provide.

TikTok also believes that Ofcom should enable VSPs to categorise and handle complaints in different ways, depending on the nature or severity of the complaint. For example, platforms should be entitled to recognise and handle vexatious complaints appropriately.

Vexatious claims and other anomalies (such as higher thresholds or more information required at some stages of the process leading to drop off in complaints) may therefore make the quantitative data requested on the number of complaints or their outcomes less valuable to Ofcom in monitoring the effectiveness of the VSP regime. We would ask Ofcom to consider alternative means of assessing the effectiveness of complaints processes.

#### Impartial dispute resolution procedure

We have carefully considered the draft guidance that relates to the obligation on VSPs to provide for an impartial out-of-court dispute resolution procedure. TikTok supports the goal of ensuring that complaints are handled in a fair and effective way and recognises that in some cases there may be an important role for impartial dispute resolution procedures. TikTok also believes that the approach must be proportionate and consistent with the legislation. TikTok also believes there should be clear guidance for both users and VSPs on the types of disputes that can be considered within the scope of the procedure, and the processes which can and should be used for those disputes.

In its current form Ofcom's proposed guidance places a considerable burden on VSPs, which we would suggest should be replaced by an approach which has proportionality as its overriding principle. In particular:

- The Guidance expects each VSP under Ofcom's jurisdiction to create their own extensive procedures, from scratch, which would be in place only for a very short period until the Online Safety Bill (which does not contain these provisions) comes into force.
- Ofcom has jurisdiction over only a small number of VSPs, which will mean that the biggest social media platforms have no obligations in this area while VSPs under Ofcom's jurisdiction would be subject to wide ranging obligations that currently go well beyond comparable regimes in EU member states. As acknowledged in A1.29, this would affect the competitive environment for VSPs and in particular reduce the

opportunity for smaller VSPs to enter the market and compete with larger and more established VSPs.

- Many more types of dispute would be subject to the procedure under the draft guidance than the legislation requires. The proposed burden on VSPs is already significant without the draft guidance adding to that burden by expanding the scope of disputes subject to the procedure.
- TikTok believes the proposed approach may not be consistent with the statement at section A1 of Ofcom's draft guidance (considering the impact of the VSP regime) which confirms that the UK adhered to the minimum requirements of the Directive.

As Ofcom is proposing to place the responsibility onto the VSPs, of which there are very few, and as even those few will not be subject to the regime for very long, it is imperative that the regime is proportionate, is defined in line with the legislation and does not place unreasonable burdens on companies. We note that several other jurisdictions are facilitating this requirement via a mechanism set up by their regulator rather than placing the requirement on individual companies.

We also note that the government's response to the implementation of the VSP elements of the Directive, published on 24 July 2019, "*expect[ed] Ofcom to consult on their approach to overseeing these requirements to ensure that the process they follow is proportionate*" [emphasis added]. This approach is consistent with that set out in the Directive; Recital 45 references the need to "*set out proportionate rules*" on VSPs and article 28b(3) requires the appropriate measures to be "*practicable and proportionate*". More generally, the government recorded its intention "*to implement the Directive in the most straightforward way possible - adhering to the minimum requirements of the Directive*" [emphasis added]."

#### *Types of complaints eligible for resolution under the dispute resolution procedure*

TikTok believes further consideration of the types of dispute eligible for resolution under the dispute resolution procedure would be appropriate. On our reading, the draft guidance goes well beyond the legislation in relation to the types of disputes that will be eligible to be considered and settled by the dispute resolution procedure. In this respect, we note that the Directive states that "*Member States shall ensure that out-of-court redress mechanisms are available for the settlement of disputes between users and VSPs*" relating to the application of the appropriate measures (those set out at Article 28b paragraph 3(a) to (j), including for example "providing for parental control systems" and "providing for effective media literacy tools"). The Act states that the impartial out-of-court procedure should be available for the resolution of disputes either relating to "*(i) the implementation of any measure set out in Schedule 15A*" (for example, mechanisms for viewers to report harmful material) or "*(ii) a decision to take, or not to take, any such measure*" [emphasis added]. The legislation therefore requires only that the procedure should be available to settle disputes that relate to whether and how VSPs have implemented the appropriate measures. Those measures relate to actions taken by VSPs to include certain

terms and conditions, to provide functionality and to establish and operate mechanisms, systems and procedures. They are not measures requiring VSPs to take particular decisions about content or to take decisions in a particular way. The Act does not, therefore, require the procedure to be available for resolution of disputes about decisions taken by VSPs in individual cases under the measures they have implemented.

The draft guidance appears to extend further than these well-defined categories and states that the procedure should broadly “allow users to challenge decisions taken by the VSP provider and seek redress”. It specifically refers to “content removal” as a type of dispute that may be handled by the dispute resolution procedure. The draft guidance does not explain in what way the Act requires the procedure to be available for resolution of that type of dispute. By way of comparison, our understanding is that Member States that have implemented or are in the process of implementing the Directive including Netherlands, Austria, France and the Czech Republic have not brought decisions on content removal within the scope of the impartial dispute resolution procedure. We do not believe that it was the intention of the Directive that content removal decisions would be within scope. We note for example Recital 48 to the Directive: *“In light of the nature of the providers’ involvement with the content provided on video-sharing platform services, the appropriate measures to protect minors and the general public should relate to the organisation of the content and not to the content as such”* [emphasis added]. As the measures themselves do not relate to the content as such then resolution of disputes should follow the same approach .

Enabling a dispute resolution procedure to be available for disputes regarding content removal would place a disproportionate burden on VSPs under Ofcom’s jurisdiction. This point can be illustrated by considering the volume of content which is available on, and may be taken down by, VSPs. In the second half of 2020 alone, TikTok removed 89,132,938 videos globally, which equates to less than 1% of the total videos uploaded. The vast majority of these videos were identified and removed before a user reported them and before they received any views. The resource required to create, operate and fund a dispute resolution structure, even for a fraction of these removals, is likely to be very significant. Consideration should be also given to the perverse incentives this might set for moderators and users, particularly if there are no disincentives for abuse of the system by bad actors (for example, if cost for dispute resolution is borne solely by VSPs).

In addition, when considering the criteria for the eligibility of disputes to be handled under this procedure, it will be important to take into account legal restrictions in sharing users’ personal data with third parties in the event that the person or body handling the dispute is a third party to the VSP. The top three reasons for content removal on TikTok in 2020 were minor safety; adult nudity and sexual activities; and illegal activities and regulated goods. Provision of information on these cases would be fraught with difficulties. Without access to the necessary data (including the identities of the people involved), which in many cases would not be appropriate or may not be lawful to share, the decision-making process would not be based upon facts and

could lead to inappropriate decisions being reached. In this respect, we note that Ofcom's draft guidance acknowledges that when considering what are the appropriate measures for implementation, VSPs must consider, as part of their proportionality analysis, the rights and legitimate interests of users including users' rights to privacy and data protection.

#### *Procedural elements of the dispute resolution procedure*

TikTok would welcome further details in the draft guidance regarding the dispute which the Act identifies, namely a dispute about whether and how VSPs have implemented the appropriate measures. In particular, TikTok would welcome guidance on the procedural elements regarding the dispute resolution procedure and the process that would be undertaken by the individual or body carrying out the dispute resolution procedure. Procedural requirements are important to prevent abuse by vexatious complainants, which would cause delay in handling genuine, urgent or serious disputes and would therefore seriously risk the overall fairness and success of any system. This would also significantly increase the financial cost of the process which may be very high and fall disproportionately on VSPs within Ofcom's jurisdiction.

It is standard practice for dispute resolution procedures in other sectors (such as aviation, travel, gambling, the media, broadband and telephone communication services, postal and storage services) to have the following features:

- The company may refer disputes to the dispute resolution procedure, in addition to the customer.
- Complainant customers may access the dispute resolution procedure only when they have fully exhausted the applicable internal complaints procedure.
- Complainants are required to identify themselves for the purposes of their application and this will be recorded on a register of applications (in accordance with data protection)
- Complainants must provide comprehensive information about their complaint upon application (often in a specified standard form).
- There is a cost to the complainant of accessing the procedure (in addition to the cost to the company with whom there is a dispute).
- Complainants are required to submit their application to the dispute resolution procedure within a certain time period, which can range from a matter of days to months depending on the industry.
- Applications to the dispute resolution procedure may be summarily rejected following initial review for being frivolous (entirely without merit) and/or vexatious.

Ofcom might consider encouraging VSPs to adopt the above features to ensure the process is effective, proportionate and fair both to users and VSPs.

TikTok would welcome confirmation from Ofcom’s final guidance that VSPs will be encouraged to adopt the above practices in any dispute resolution procedure established by VSPs and to have flexibility in adopting such practices. We consider it important that a VSP has a measure of flexibility in adopting these practices, for example by determining the period of time for filing of a complaint (and when time starts running for submission), the steps a complainant must take to participate in the process and appropriate arrangements for the fair sharing of associated costs.

#### *Outcomes of the dispute resolution procedure*

The draft guidance states that dispute resolution procedures could bring about a range of outcomes outlined in Paragraph 4.132, including *“the removal of content; sanctions against offending users; the reversal of wrongful content removal or sanctions; issuing an apology.”* TikTok’s view is that these outcomes are not appropriate for an impartial decision-maker and alternative outcomes should be recommended or this matter left to the VSP to determine. It is unclear to us how an independent third party issuing its decision on a dispute may be able to deliver such outcomes. TikTok’s view is that more appropriate outcomes would be to publish the adjudication or to require the VSP reconsider the complaint. We would welcome a commitment from Ofcom to consult further on the application of the outcomes given the very extensive powers Ofcom envisages the person or body responsible for the independent dispute resolution will have.

Another possible outcome noted in the draft guidance is *“changes to processes or policies”*. Whilst we agree with the importance of continually improving processes and policies, again it is unclear how this outcome could be achieved if the procedure was handled by a fully independent third party rather than a designated person or department within the VSP. We would also note that specific complaints are unlikely to provide the relevant third party arbiter with a sufficiently comprehensive understanding of the underlying process or policy, which may serve much broader purposes and be subject to a wide variety of additional considerations beyond the confines and/or context of the dispute in issue. TikTok would welcome further clarity from Ofcom as to how these feedback loops can effectively operate without Ofcom considering that the impartiality of the procedure has been compromised.

### **Media Literacy Tools and Information**

*11. Do you have any comments on the draft guidance about media literacy tools and information?*

We agree that media literacy is critical to empower users with the skills, tools and information they need to critically assess information and know how to proactively and visibly identify and respond to information. These education and resilience based initiatives are a key part of a

comprehensive and inclusive strategy that empowers users of all ages to navigate the online world safely and securely.

TikTok is therefore encouraged by Ofcom's focus on media literacy in the draft guidance. As outlined earlier in this response, and in our September 2020 submission, TikTok has taken a range of actions to improve users' media literacy and to raise awareness of the tools and information available to them.

As the draft guidance highlights, this is an evolving area of work and there is an important role for partnerships and sharing best practice. Ofcom can play an important facilitation role in helping achieve this. In doing so it should be recognised that there are potential legal and privacy restrictions on what quantitative data can be collected or shared on users, tools and campaigns.

We welcome Ofcom's acknowledgement that VSPs should consider the type of information or tools needed based on the nature of their service and types of users on it. Each platform is unique in its structure and user base, and content and campaigns that are appropriate for some will not be for others.

TikTok is a full-screen, video-first VSP. Content that performs well on TikTok needs to be creative, authentic and engaging. This applies to media literacy campaigns, just as it does to user-generated content. Campaigns or content designed for other platforms that are text-based or static will not be as effective as TikTok-native interventions.

We therefore believe it is best for VSPs to decide how best to achieve their media literacy obligations, taking into account these nuances and reflecting that it is the platform that will have both the detailed understanding of their userbase's needs and requirements, applying their specific knowledge and expertise to design effective on-platform solutions.

## **Proportionality**

*12. Do you have any comments on the draft guidance provided about the practicable and proportionate criteria VSP providers must have regard to when determining which measures are appropriate to take to protect users from harm?*

TikTok welcomes guidance concerning the practicable and proportionate criteria set out in Section 5 of the draft guidance and generally we agree that it is for VSPs to decide which measures should be implemented, according to the principles set out.

Proportionality is the key criteria - the steps taken by each VSP to protect their users will naturally need to be different depending on its size, nature of a service and the sophistication of each platform. Factors such as the volume of content available are relevant not only to

determining the risk of harmful material being available but also to the question of whether particular measures would be practicable and effective. We welcome confirmation from Ofcom that when considering whether it is proportionate to implement a certain measure, VSPs may take into account “any other measures which have been taken or are to be taken”: we encourage Ofcom to consider VSPs’ implementation of measures in the round and recognise that if VSPs have obtained reliable information indicating that certain measures are not likely to be effective in protecting users from harmful material and/or restricted material that they will not be expected to implement these.

Part of considering proportionality must involve taking into account what is in the power or control of VSPs. Paragraph 5.11 of the guidance states providers should be aware of the reach of their service beyond the boundaries of their own platform, with content potentially being re-broadcast across other mediums.

In relation to the suggestion that VSPs should consider the reach of the content shared on their platform: while there are some restrictions on sharing TikTok content (for example, content created by a user under 16 cannot be downloaded) it is important to note that VSPs have no control on the onward dissemination or use of content that has been uploaded to other platforms.

We agree that it is important VSPs take into account the impact the measures that are implemented may have on the general public and the importance of ensuring that content that may initially seem harmful but is in the public interest is permitted. This is one of the reasons why we encourage Ofcom to not be overly prescriptive in its guidance about the types of content that may constitute “restricted material”.

In relation to the “Additional steps to protect users”, we agree that the steps listed (embedding a safety-first approach, external engagement, etc) may be very valuable to VSPs and TikTok has already embedded many of these. However, Ofcom has noted that “These are related to the protection measures but are not necessarily a requirement of the VSP Framework” [emphasis added]. As these steps are not required by the legislation, we would be grateful if Ofcom would confirm that these steps are illustrative only and that VSPs will not be expected to implement these in order to comply with the Act.

## **Assessing and Managing Risk**

*13. Do you have any comments on the draft guidance about assessing and managing risk?*

Assessing and managing risk is core to the approach taken by TikTok’s Trust and Safety team. Due to the unique nature of each platform and the risks that may arise we believe it will be most effective for VSPs to determine the best method of assessing and managing these areas, in discussion with the regulator.

We involve a number of internal and external experts to help inform how we assess and manage risk, particularly through our European Safety Advisory Council. The Council enables us to bring different, fresh perspectives on the challenges we face with subject matter expertise provided on our content moderation policies and practices. This enables not only to develop forward-looking policies that address the challenges we face today, but also to help us to identify emerging issues that affect TikTok and our community in the future.

We welcome Ofcom's observation in Paragraph 6.46 that numerical indicators should be put into context. TikTok is committed to transparency, whether through our regular transparency reporting (<https://www.tiktok.com/safety/resources/transparency-report-2020-2>) or Transparency and Accountability Centre outlined elsewhere in this submission, however context on this information is crucial. An increased rate of content removal does not necessarily mean a higher level of harm - this could imply several things including better enforcement action or improved education and use of reporting tools.

## Costs

14. *Do you have any comments on the impact assessment in Annex 1, including the potential impacts to VSPs outlined in tables 1 and 2, and any of the potential costs incurred (including any we have not identified)?*
15. *Do you have any comments on our provisional assessment that the potential costs for providers are proportionate to achieve the regulatory requirements of the regime?*

In our response to question 10, we have confirmed that we believe that Ofcom's guidance on the impartial out-of-court dispute resolution procedure needs to be more proportionate and consistent with the legislation. The current draft guidance presents a risk of imposing on VSPs a requirement which is burdensome, open to abuse and for these reasons, potentially very costly. The assessment of potential costs does not appear to have sufficiently taken this into account.

Similarly the assessment of impact on stakeholders (from A1.18) does not take into account that a dispute resolution procedure that lacks clarity, fails to deter vexatious complaints and could be easily overwhelmed would have a detrimental impact on users with legitimate complaints. Ofcom's proposed guidance relating to the dispute resolution procedure also does not accord with the statement at A1.31 that proportionality is an integral part of the guidance and that the UK adhered to the minimum requirements of the Directive.

## Other

16. *Do you have any comments on any other part of the draft guidance?*





With regards to other elements of the guidance, Paragraph 7.16 outlines the ability for users to address complaints directly to Ofcom. We welcome that the draft guidance requests users to first raise their complaints with VSPs directly, and would encourage steps in the process to incentivise this.

We also would welcome clarity on some specific areas of the regulation with regards to implementation. In particular we would welcome clarity on intended consultation periods for updated guidance in the future, and on the proposed approach to regulatory enforcement in the interim period until all guidance, including on advertising related elements, is finalised.

More broadly we would request Ofcom share any trends emerging from direct complaints, to notify VSPs and allow for sufficient time to inform of action already taken, or address issues that have not been raised with the platform, prior to any enforcement action being discussed.