

Written evidence submitted by TikTok (SMH0081)

Chi Onwurah MP
Science, Innovation and Technology Committee
House of Commons
Palace of Westminster
London
SW1A 0AA

10/04/2025

RE: Follow-ups from 25 February oral evidence session

Dear Chair,

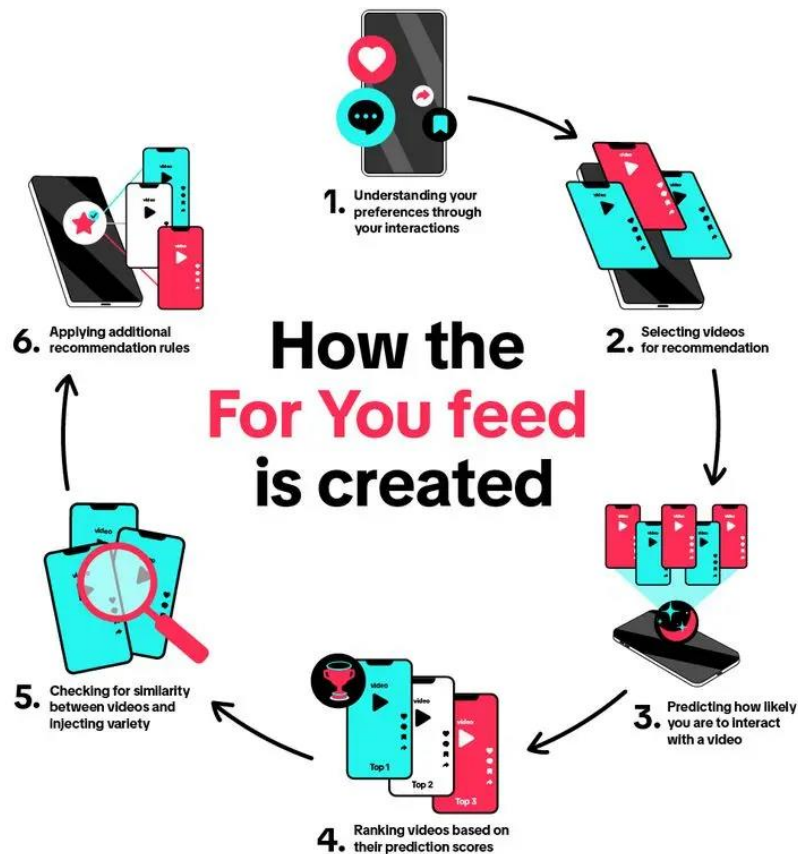
Thank you for your letter of 20th March 2025, and for the opportunity to give evidence to the Committee's inquiry on social media, misinformation and algorithms. Providing a safe and secure environment for all those who use our platform is a core part of our mission and we consider constructive engagement with your Committee and all those who want to see a safe internet to be a vital part of our pursuit of this goal.

In our written and oral evidence, we outlined how our personalised recommendation system and the For You Feed (FYF) works. The algorithmic personalisation system understands a user's interactions with certain pieces of content - for example, if they watch the entire video, like the video, or say they are not interested. It then selects videos for recommendation and, using that understanding of the user's interaction with previous pieces of content, predicts how likely they are to interact with a video. It then ranks the videos based on the prediction, checks for similarity between videos and injects some variety before finally serving these on the user's FYF.

This recommendation system works in tandem with TikTok's trust and safety algorithms and teams. We maintain strict content eligibility standards for the FYF that prioritise safety and diversity. While we remove content which violates our community guidelines, we may also make other content ineligible for the FYF, or restricted to be only available to users over 18, when it is not appropriate for a broad audience. This includes:

- content showing activity that involves visible or imminent moderate physical harm
- content showing graphic or potentially distressing footage of events that are in the public interest to view, such as clashes with law enforcement or the aftermath of a bombing or natural disaster
- conspiracy theories that are unfounded and claim that certain events or situations are carried out by covert or powerful groups, such as "the government" or a "secret society"
- repurposed media, such as showing a crowd at a music concert and suggesting it is a political protest
- unverified claims related to an emergency or unfolding event

- potential significant high-harm misinformation while it is undergoing a fact-checking review



We have invested heavily to provide a variety of resources to help our users understand how the FYF works. In 2022, we rolled out a series of tools to allow users to understand and control what type of content is recommended to them. This includes our 'Why this video' feature, which summarises why a certain piece of content has been recommended to them, for example if it is a popular video in the region they are based. We have also introduced tools to indicate that a user might not be interested in certain types of content, to filter out keywords from the FYF, or to refresh their FYF entirely and experience the platform as if they are a new user.

At TikTok, we share the Committee's view that moderation is important to ensuring that harm is minimised and that social media platforms are as safe as possible for their users. We want to ensure that we create a safe, trustworthy environment in which users can create, discover and enjoy content.

As discussed during the oral evidence session, TikTok has strict Community Guidelines that sit at the heart of our efforts to protect our community. These Community Guidelines contain specific prohibitions on harmful misinformation, hateful content, and undisclosed synthetic or manipulated media. We regularly update these policies to keep pace with new or emerging threats to our users' safety, and invest heavily in AI and human moderation to consistently and swiftly enforce content

which may violate these policies.

In the aftermath of the attack in Southport on 29 July, TikTok swiftly launched a command centre in line with our crisis management protocols. This command centre brought together over 100 people across 10 teams across the world, helping direct moderators who work 24/7 to ensure we were responding swiftly to content which may have violated our Community Guidelines. This work built on TikTok's safety-by-design approach, including ensuring that all content uploaded to TikTok is reviewed by our AI systems, and escalated for human review if potential issues are identified.

During the peak of the unrest, these teams oversaw the removal of tens of thousands of videos and comments for violations of our Community Guidelines concerning harmful misinformation and hateful conduct.

We regularly publish information on how proactive and quick we are in enforcing our Community Guidelines. For Q3 2024 98.2% of content removed by TikTok for violating our Community Guidelines was done proactively - i.e. detected by our teams, not in response to a user report or external escalation. 92.1% of this content was removed within 24 hours of being uploaded. Furthermore, 85.5% of this content was removed before it had received a single view, and 98.2% before it had received 1,000 views. For content reported to us by a user, 82.2% of reports were responded to in under 2 hours, and 99.7% within 24 hours.

Our work on enforcement of these Community Guidelines was complemented by a wider set of actions including engagement with Government, law enforcement authorities and community groups. To ensure users were directed to appropriate information and context around the unrest we also implemented search and wider media literacy interventions. These interventions warned of the risks of mis and disinformation, leading users to an in-app hub with tips on how to assess whether content is authoritative, as well as educational videos from our external fact-checking partner Logically Facts.

TikTok's priority is always to ensure that we are providing a safe and enjoyable environment for our users. We are happy to engage with the Committee further on how we develop, implement and enforce our safety policies in more detail. As mentioned in our oral evidence, the Committee has a standing invitation to visit our Transparency and Accountability Centres that would provide this deep dive into how we keep our community safe, as well as the TikTok algorithm and recommendation system.

The Transparency and Accountability Centre builds on TikTok's wider transparency efforts, including our Research API, and the role of US company Oracle which is responsible for line-by-line code reviews and for auditing any updates before uploading to the Apple and Google App Stores.

As I'm sure the Committee understands, there is some information regarding our trust and safety measures and the workings of our algorithm that cannot be disclosed publicly. This can not only be for reasons of commercial confidentiality, but also as the disclosure of some information can undermine the efficacy of our trust and safety systems, increasing the risk that malicious actors exploit our service and bypass the very same protections the Committee agree are so vitally important to keeping our users safe.

Once again, we would like to thank the Committee for the opportunity to provide written and oral

evidence, and hope the information provided has helped build its understanding of how TikTok seeks to protect its users, particularly in times of crisis. We look forward to working closely with you and the Committee in future on these important issues.

Yours sincerely,

Ali Law

Director, Public Policy and Government Relations
TikTok UK and Ireland

10 April 2025