

Consultation response form

Please complete this form in full and return to vspreulation@ofcom.org.uk

Consultation title	Guidance for VSP providers on measures to protect users from harmful material
Full name	Megan Manson
Contact phone number	020 7404 3126
Representing (delete as appropriate)	Organisation
Organisation name	National Secular Society
Email address	megan.manson@secularism.org.uk

Confidentiality

We ask for your contact details along with your response so that we can engage with you on this consultation. For further information about how Ofcom handles your personal information and your corresponding rights, see [Ofcom's General Privacy Statement](#).

Your details: We will keep your contact number and email address confidential. Is there anything else you want to keep confidential? Delete as appropriate.	Nothing
Your response: Please indicate how much of your response you want to keep confidential. Delete as appropriate.	None
For confidential responses, can Ofcom publish a reference to the contents of your response?	

Please complete this form in full and return to vspreulation@ofcom.org.uk

Your response

Introduction

This submission is made by the National Secular Society (NSS). The NSS is a not-for-profit, non-governmental organisation founded in 1866, funded by its members and by donations. We advocate for separation of religion and state and promote secularism as the best means of creating a society in which people of all religions and none can live together fairly and cohesively. We seek a diverse society where all are free to practise their faith, change it, or to have no faith at all. We uphold the universality of individual human rights, which should never be overridden on the grounds of religion, tradition or culture. We promote free speech as a positive value.

We welcome this opportunity to respond to Ofcom's consultation on its video-sharing platform (VSP) guidance.

VSPs are not merely a source of recreation and entertainment – they have become a powerful tool for information-sharing and awareness-building world over. Critically, they empower ordinary people to express themselves and share their stories in a way that they cannot via more traditional forms of media such as broadcast media or newspapers. They are particularly crucial for oppressed and marginalised groups, who often lack any other means to speak out about their oppression and seek justice.

For this reason, it is essentially that VSPs enable free speech – including speech about contentious and controversial ideas that many people may find 'offensive' or even 'hateful'. We are concerned that as it stands, Ofcom's guidance does not provide adequate free speech protections and will prompt VSPs to over-censor content. We have outlined our concerns in our answers below. We have not answered questions that are beyond our remit as an organisation.

Question 1: Do you have any comments on Section 3 of the draft guidance on harmful material and related definitions?

We are very concerned that Ofcom's definition of "relevant harmful material" is too vague and that Section 3 does not provide adequate guidance for protecting free speech.

Although the consultation document on the draft guidance makes many references to Ofcom's commendable support for freedom of expression, the guidance itself places less emphasis on this and no duty at all on VSPs to protect free speech. We fear that this, combined with the broad and vague definitions for material 'likely to incite hatred', will result in VSPs being disincentivised to protect free speech, and instead motivated to apply excessive moderation.

It is completely reasonable for Ofcom to require VSPs to moderate restricted material, in addition to removing material which would be a criminal offence. It is equally reasonable for Ofcom to require VSPs to protect users from material "likely to incite violence", because most violent acts would also be a criminal offence.

However, problems arise from Ofcom's requirement for VSPs to protect users from any material likely to incite "hatred against a group of persons or a member of a group of persons based on particular grounds". These grounds are defined in the guidance as "sex, race, colour, ethnic or social

origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age, sexual orientation”.

There is no definition provided in the guidance itself as to what ‘incitement to hatred’ means aside from that it “should be understood as having its usual meaning in everyday language”. This is extremely vague, because the understanding of “incitement to hatred” has no one “usual meaning”. ‘Hatred’ is largely subjective and different individuals will have different interpretations as to what constitutes “incitement to hatred”.

These concerns are shared by VSPs. As Ofcom mentions in its consultation guidance:

“One platform had concerns that a lack of clarity about what is/is not acceptable content moderation could lead to providers going over what is required, which could have a negative impact on freedom of expression. Another platform said that while it has an appeal process in place for moderation decisions, to facilitate effective moderation and maintain freedom of expression, legislation needed to provide clear, concise and robust definitions of the terms they are required to enforce.” (Ofcom, *Video-sharing platform guidance Consultation on guidance for VSP providers on measures to protect users from harmful material*, p. 22 para 4.19)

There are many videos hosted by VSPs that strongly criticise or ridicule religion. These include videos exposing harms perpetuated by religious practices, refuting claims made in religious texts, encouraging people to leave a religion, or simply making fun of religious beliefs that the video creator finds ridiculous. To religious people, such videos could be interpreted as ‘inciting hatred’ and could therefore prompt calls for them to be removed. But in a free and democratic society, all must be free to criticise or ridicule any ideas or concepts – including religious ones.

There have already been many cases of religious groups weaponising claims of ‘incitement to hatred’ in an attempt to censor videos that criticise their religion or that they consider ‘blasphemous’. These have frequently resulted in videos being removed. In some cases, religious organisations rally support from devotees to encourage mass reporting of ‘offensive’ content. Doing so increases the likelihood of the content being removed – not because the material is genuinely harmful, but because mass reporting is more likely to trigger automated processes that result in content removal, and puts human moderators under greater pressure to censor the content.

One example is the comedian Waleed Wain (a.k.a Veedu Vidz) who makes videos satirising Islamist extremism and anti-Muslim bigotry. His Facebook page has been removed on multiple occasions because some users repeatedly reported his videos for “breaking the Facebook community guidelines.” In one incident in February 2018, the Veedu Vidz Facebook page shared a video entitled “Halal Movie Review: The Lion King”, which made fun of a well-known Islamist preacher. Within 24 hours of sharing the video, the Veedu Vidz page was “unpublished for repeatedly posting things that don't comply with the Facebook terms”. Further information on this case can be found here: <https://www.secularism.org.uk/news/2018/02/nss-criticises-facebook-over-ban-on-halal-lion-king-satirist>

Following widespread criticism from individuals and organisations including the NSS, Facebook reversed its decision. However, the fact that the page was banned at all demonstrates how easily VSP guidelines regarding ‘incitement to hatred’ can be weaponised by those wishing to censor content that ridicules religion.

Another example is an attempt made by the Christian Congregation of Jehovah’s Witnesses (CCJW) to silence Lloyd Evans, a former Jehovah’s Witness elder and critic of the religion. In 2020, Evans was selected as a core participant in the Independent Inquiry into Child Sex Abuse (IICSA) to give evidence regarding child abuse in the Jehovah’s Witness religion. In an attempt to have Evans

removed as a participant, CCJW sent IICSA a dossier of what they called 'hate speech' compiled from videos and user comments on Evans' YouTube channel, which features videos including fair and legitimate criticisms of the CCJW. Thankfully IICSA did not accept CCJW's claims. Further information, including links to the relevant documents, can be found on Lloyd Evans' YouTube series of videos about the case: <https://youtu.be/D596VuaqSrM>

Those who criticise or ridicule religion are not the only content producers at risk of censorship under 'incitement to hatred' rules. Religious groups themselves are equally at risk. Many religious groups have very conservative views about same-sex relationships or the role of women in society, for example. For many women and LGBT+ people, these views can be highly offensive and distressing, and easily interpreted as 'hateful'. Religious groups and individuals expressing such views on VSPs may also find themselves the target of mass reporting for 'incitement to hatred'. But to censor religious groups for simply expressing these teachings alone risks imposing too great a restriction on freedom of expression and freedom of religion or belief. Additionally, it is arguably in the interest of the wider public for such videos to be uncensored, as it can assist individuals and organisations who wish to avoid dealings with groups who hold such views.

Ofcom's guidance refers to ECHR case law on freedom of expression as a guide. However, VSPs are unlikely to have the time and resources necessary to consult this case law in depth in every case where they need to decide whether particular content 'incites hatred'. Because the penalties for permitting content that Ofcom deems harmful are severe (including financial penalties of up to £25,000 or suspending VSPs outright), VSPs are likely to err on the side of censorship in order to 'be on the safe side', as there are no equivalent penalties for failing to protect free speech.

The guidance also references the Alan Turing Institute report *Understanding online hate: VSP Regulation and the broader context*. We have some concerns about the use of this report as a guide for dealing with hateful content, as some of the descriptions of 'hate' described in the report are again vague:

"Overt forms of hate are usually more aggressive and can include 'amplifying' elements, such as swear words. Covert forms are harder to identify and may intentionally be expressed in a misleading way through 'obfuscation'. Such content will be harder for content moderation systems to detect. Some forms of hate will be difficult even for trained experts to recognise due to the use of 'code words' and complex forms of language."
(p.3-4)

While it is true that certain hate groups do use in-group 'code words' and some individuals may employ obfuscation techniques to avoid detection, discretion is needed to ensure over-censorship does not occur through misplaced accusations of 'dog whistle' hate. It may inadvertently result in moderators (either human or automated) searching for hate that is not there.

Finally, we are concerned about the inclusion of "political or any other opinion" among the grounds that a video may be considered to 'likely to incite hatred'. As with religion or belief, we are concerned that moderators will be unable to effectively distinguish between robust criticism of a "political or any other opinion", and incitement to hatred of those who may hold such opinions or beliefs. Again, there is no adequate explanation for this included in the guidance.

Our recommendations: We strongly recommend that the guidance includes clearer instruction on the duty of VSPs to have due regard for freedom of speech and freedom of religion or belief, together with more comprehensive explanations about what does *not* constitute material that is considered "likely to incite hatred" with regard to religion or belief.

For example, the Racial and Religious Hatred Act 2006 explicitly protects "discussion, criticism or expressions of antipathy, dislike, ridicule, insult or abuse of particular religions or the beliefs or practices of their adherents". Similar wording included in Section 3's subsection on "Material likely

to incite violence or hatred” would help add clarity and protect freedom of speech around religion or belief.

Question 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?

Please see our answer to Question 1. As Ofcom mentions in its consultation guidance regarding terms and conditions which prohibit relevant harmful material:

“One platform had concerns that a lack of clarity about what is/is not acceptable content moderation could lead to providers going over what is required, which could have a negative impact on freedom of expression. Another platform said that while it has an appeal process in place for moderation decisions, to facilitate effective moderation and maintain freedom of expression, legislation needed to provide clear, concise and robust definitions of the terms they are required to enforce.” (Ofcom, *Video-sharing platform guidance Consultation on guidance for VSP providers on measures to protect users from harmful material*, p. 22 para 4.19)

We also note that one VSP expressed a casual disregard for freedom of expression regarding the enforcement of terms and conditions:

“Two platforms responded that balancing freedom of expression with safety was a top priority when moderating content and drafting their terms of service. However, another respondent recognised that online platforms have no obligation to promote freedom of expression. This platform said that its community guidelines were very clear about what content is permitted on the platform and that anyone found in violation of the guidelines could not make a “freedom of expression” argument to prevent the platform from taking appropriate action.”
(Ofcom, *Video-sharing platform guidance Consultation on guidance for VSP providers on measures to protect users from harmful material*, p. 21 para 4.18)

The fact that some VSPs place little or no importance on freedom of expression is all the more reason for Ofcom to include clearer statements about what does not constitute material that is “likely to incite hatred” (e.g. videos that robustly criticise or ridicule religion).

Finally, we welcome the statement at 4.51 in the guidance:

“As noted in Section 3, material which might impair the physical, mental or moral development of under-18s is likely to evolve over time and VSP providers should ensure they remain informed about changing attitudes.”

This is a welcome acknowledgement that societal attitudes constantly change over time, which is all the more reason to ensure material is not erroneously censored for being ‘likely to incite hatred’.

Question 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?

We welcome and agree with Ofcom’s statement at 4.98 of the consultation document:

“It is important that all users of all VSPs have an opportunity to challenge decisions that have been made by platforms. It is fundamental to ensure that user’s rights and legitimate interests are protected and to mitigate against over-takedown of content.”

However, we think that such a provision will be of little use if there are not stronger protections for free speech in the guidance regarding material that is “likely to incite hatred”.

Question 12: Do you have any comments on the with 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?

We welcome and agree with Ofcom’s statement in the guidance at 5.35:

“In designing and implementing protection measures, VSP providers should also take into account the impact such measures may have on the general public. For example, some content which might initially seem harmful, may actually be in the public interest. Videos containing news content are likely to fall within considerations of general public interest and in Section 4 we suggest ensuring that robust dispute resolution processes are in place which give careful consideration to this content.”

We agree that content which “might initially seem harmful, may actually be in the public interest”. This will include material that may initially seem “likely to incite hatred”, such as material that robustly criticises religion. However, more explicit protections for such material need to be in place in Section 3 to ensure much material that is in the public interest is not censored.

Question 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)?

We are disappointed that the risks to freedom of expression were not explicitly acknowledged in the impact assessment for viewers or creators. This should have been identified as a key issue for Ofcom to address and bear in mind in its impact assessment.