making communications work
for everyone

Video-sharing platform guidance

Guidance for providers on measures to protect users from harmful material

Guidance for providers on measures to protect users from harmful material – Welsh overview

Publication date: 24 March 2021

Contents

Section

1. Overview

What this guidance covers

New legislation applying to UK-established video-sharing platform (VSP) services came into force on 1 November 2020.

This guidance is intended to support VSP providers in understanding their statutory obligations under the new regime, referred to as "the VSP Framework". These obligations include requirements to take measures to protect users from harmful material in videos. This document:

- provides a background to the rules and requirements of the VSP Framework;
- explains what types of content might constitute harmful material in videos;
- provides an explanation of the measures platforms can take to protect users from that harmful material, as well as guidance on how to implement those measures effectively;
- explains the practicable and proportionate criteria providers must consider when deciding which measures to take, which include considerations about the individual characteristics of VSPs and the rights and legitimate interests of users;
- encourages the use of additional steps to help protect users, including undertaking risk assessments to help inform which measures they take and how those measures are implemented; and
- provides information about Ofcom's approach to monitoring and enforcement.

Overview of the requirements

The VSP Framework requires providers to take appropriate measures to protect:

a) the general public from **"relevant harmful material".** This includes:

   i) incitement to violence or hatred against particular groups

   ii) content which would be considered a criminal offence under laws relating to terrorism; child sexual abuse material; and racism and xenophobia

b) under-18s from **"restricted material".** This includes:

   i) material which has, or would likely be given, an R18 certificate

   ii) material which has been deemed, or would likely be deemed, unsuitable for classification (such as sadistic violence or torture)

   iii) other material which might impair the physical, mental or moral development of under-18s

We refer to these two categories of material as "harmful material" and discuss the two requirements collectively as the "requirement to protect users from harmful material". There are also other administrative requirements for VSPs which we outline in Section 2, and specific requirements relating to advertising standards which will be dealt with in a separate publication.

Providers can choose to take measures set out in the VSP Framework to protect users from harmful material. These are summarised as follows:

- measures relating to terms and conditions
- measures relating to the reporting, flagging or rating of content
- access control measures such as age assurance and parental controls
- complaints processes (and a separate requirement to provide for an impartial procedure for the resolution of disputes)
- media literacy tools and information

These measures must be implemented in a way that protects users from harmful material. VSP providers are required to determine whether it is appropriate to take a particular measure, according to whether it is practicable and proportionate to do so. A set of considerations are included in the VSP Framework including the size and nature of the platform; the type of material on the platform and the harm it might cause; and the rights and legitimate interests of users.

Nature of this guidance

The VSP Framework affords VSP providers flexibility in how they protect users. This reflects the diversity of the sector and the importance of technological innovation. As such, this document is not a set of compulsory steps, but is intended to help guide providers in deciding how best to comply with the statutory requirements. It is for providers to determine which measures are appropriate for their service, but where we think effective protection of users is unlikely to be achieved without a specific approach, we say so.

Below we have categorised some of the different expectations on providers.

What are VSP providers required to do?
To protect the general public from relevant harmful material
To protect under-18s from restricted material
To provide an impartial dispute resolution procedure
Which measures are considered central to protecting users?
Having and effectively implementing terms and conditions which prohibit the uploading of relevant harmful material
Where a provider has terms and conditions requiring uploaders to notify them if a video contains restricted material, taking action in response to this notification which protects under-18s from that material
Having and effectively implementing a form of flagging and reporting mechanism
Effectively implementing robust age verification for VSPs which specialise in, or have a high prevalence of, pornography or material unsuitable for classification.

What does Ofcom strongly encourage providers to do to support compliance?
Conduct a risk management process to support decisions about which measures are appropriate for protecting users on the platform and how to implement them.
Collect information to measure the effectiveness of the measures on the platform

Ofcom's approach to assessing compliance with the VSP Framework

It is for providers to decide how they meet the requirement to protect users from harmful material. Ofcom has a duty to take steps to ensure that VSP providers comply with these requirements, which we will carry out through monitoring and enforcement.

While we acknowledge that harmful material may not be completely eradicated from a platform, we expect providers to make meaningful efforts to prevent users from encountering it. Evidence of a prevalence of harmful material appearing on a platform may require closer scrutiny. Ofcom will want to understand the measures a platform has in place, their effectiveness at protecting users and any processes which have informed a provider's decisions.

Along with engagement with providers themselves, we expect to inform our understanding of whether users are being effectively protected by monitoring complaints and engaging with interested parties such as charities, NGOs and tech safety groups. This engagement will play an important part in supporting Ofcom's decisions about areas of focus as we move through the VSP Regime.

Ofcom's understanding of which measures are appropriate and what constitutes effective protection for users will develop over time and we recognise that both platforms themselves, and the risk of harm on those platforms, will evolve. We will work with providers to ensure services are clear about what is expected of them. To support this, we will publish annual reports and we will update this Guidance where necessary.

Where Ofcom has concerns, we will act in accordance with our enforcement guidelines. Where appropriate, we will generally attempt to work with providers informally to try to resolve those concerns. Where serious concerns arise, we have the power to take swift and robust enforcement action, which may include sanctions.

The VSP Regime does not set standards for the content of individual videos. Therefore, Ofcom will not usually review individual pieces of content when investigating matters of compliance. However, in situations where more intrusive regulatory interventions are required, such as directions to remove pieces of content or the suspension or restriction of a service, Ofcom will always have regard to the user's right to freedom of expression and the right to receive information.

2. Introduction

2.1 Video-sharing platforms, or "VSPs", are subject to new regulations in the UK. In this section we provide an overview of the new regulatory framework, including Ofcom's role as the appointed regulator. We outline the obligations on VSP providers to take measures to protect users from harmful material. We also set out the scope of this guidance and how we intend it to be used by VSP providers.

Regulatory background

Ofcom is the regulator for the UK's communications services

2.2 We regulate the telecoms, broadcasting, video-on-demand, and postal industries, as well as managing civilian use of the radio spectrum. Our remit was expanded in 2020 to cover UK-established VSPs and will be expanded further in future following Government's announcement that Ofcom will be appointed as the online harms regulator.

2.3 Our principal duty is to further the interests of citizens in relation to communications matters and to further the interests of consumers in relevant markets, where appropriate by promoting competition.

2.4 Our statutory duties are set by Parliament, although we are independent from Government and funded by fees paid to us by the companies we regulate.

VSPs have been added as a new category of regulated service

2.5 The statutory framework for the regulation of VSPs is set out in Part 4B of the Communications Act 2003 ("the Act").[1] Part 4B was introduced under regulations made by the Secretary of State to implement the revised Audiovisual Media Services Directive ("AVMSD" or "the Directive") and came into effect on 1 November 2020.[2] In this Guidance we refer to the regulatory framework set out in Part 4B of the Act as "the VSP Framework" or "the VSP Regime".

2.6 The VSP Framework sets out to protect users of VSP services from harms that may result from viewing specific categories of material. In particular, to protect under-18s from potentially harmful content and to protect the general public from incitement to hatred or violence and other specific material the inclusion of which would be a criminal offence. VSPs are also required to ensure certain standards around advertising are met. The VSP Framework sets out a list of measures in Schedule 15A of the Act, which providers must consider taking, as appropriate, to secure the required protections. We outline the new requirements, including the definitions of harmful and criminal content, in further detail below.

[1] The Audiovisual Media Services Regulations 2020
[2] Some aspects of the regime, such as the requirement to notify Ofcom and to pay a fee, come into force at later dates.

Services must meet the definition of a VSP and be established in the UK to be in scope

2.7 Full details on the statutory criteria and their application for determining scope and jurisdiction have been set out in a separate publication by Ofcom, which can be accessed here. However, we set out some brief points on VSP definition and UK jurisdiction below.

VSP definition

2.8 Under section 368S of the Act, a service, or a dissociable section of a service, is a VSP if it meets the conditions listed in paragraph 2.9 below and either of the following apply:

a) the provision of videos to members of the public is the principal purpose of the service or of the dissociable section of the service;

b) the provision of videos to members of the public is an essential functionality of the service.

2.9 The additional conditions that must be met in relation to the service or dissociable section of the service are:

a) it is provided by means of an electronic communications network;

b) it is provided on a commercial basis;

c) the person providing it does not have general control over what videos are available on it, but does have general control over the manner in which videos are organised on it (which includes being organised automatically or by way of algorithms, in particular by displaying, tagging and sequencing); and

d) that person has the required connection with the United Kingdom.

2.10 The criteria set out in the Act must all be met for the definition to apply. In other words, at least one of the two criteria in paragraph 2.8 must apply and all of the conditions in paragraph 2.9 must be met for a service or a dissociable section of a service to be a VSP.

UK jurisdiction

2.11 A VSP provider will be deemed to be within UK jurisdiction if it provides the service, or a dissociable section of the service, using a fixed establishment in the UK for an indefinite period and effectively pursues an economic activity in doing so ('the case A criteria').

2.12 Where a provider is not providing the service using a fixed establishment in the UK and effectively providing an economic activity in doing so, it may still be within UK jurisdiction where it has a group undertaking established in the UK,[3] and it does not fall under the jurisdiction of an EEA State for the purposes of the AVMSD ('the case B criteria').

[3] The term group undertaking has the meaning given to it in section 1161 of the Companies Act 2006(5), except that it also includes all other undertakings having economic and legal organisational links to a VSP provider.

Ofcom was given new duties and powers under the VSP Framework

2.13 Ofcom is required to take such steps as necessary to secure compliance by VSP providers with their obligations under the VSP Framework.[4] Ofcom is also required to draw up and publish guidance concerning the measures in Schedule 15A which may be appropriate for VSP providers to take to protect users from harmful material, and the implementation of such measures. That guidance is set out in this document.

2.14 Information gathering powers enable Ofcom to demand relevant information for certain specified purposes.[5] These include assessing and monitoring compliance by VSP providers, conducting investigations into suspected contraventions of the VSP requirements and producing and publishing reports about compliance with the regime.[6]

2.15 Ofcom has the power to take enforcement action, including the power to give enforcement notifications (which may set out the steps required to remedy a contravention)[7] and to impose financial penalties of up to £250,000 or 10% of qualifying revenue, whichever is greater.[8] In certain circumstances, Ofcom may also suspend and/or restrict a service.[9] Ofcom's enforcement of the VSP Framework will be in line with Ofcom's Enforcement Guidelines.[10]

2.16 The VSP Regime is not a broadcast-style content regime and Ofcom will not resolve individual complaints about items of content. Given the volume of user-generated content, regulation is focused on the measures providers take to protect their users from harmful content.

2.17 This Guidance is drafted in light of the Human Rights Act 1998 and the European Convention on Human Rights ("the Convention"). In particular, the right to freedom of expression, as expressed in Article 10 of the Convention, which includes the right to hold opinions and to receive information and ideas without interference by public authority. Such freedoms can be subject to restrictions if they are prescribed by law and are necessary in a democratic society in pursuance of a legitimate aim.

2.18 In deciding whether a measure is appropriate, Ofcom will take into account the rights and legitimate interests at stake, including service providers and users who create, upload or view material, as well as the general public interest.

2.19 In situations where more intrusive regulatory interventions are required, such as directions to remove pieces of content or the suspension or restriction of a service, Ofcom will always

[4] Section 368X of the Act
[5] Section 368Z10 of the Act
[6] Section 368Z11 of the Act
[7] Sections 368Z2 and 368Z3 of the Act
[8] Section 368Z4 of the Act
[9] Sections 368Z5 and 368Z6 of the Act
[10] These Enforcement Guidelines are due to be updated to reflect Ofcom's new powers and a consultation on these changes will be conducted in 2021. Until then, we believe the Enforcement Guidelines provide an appropriate and applicable framework for investigations into compliance with the VSP requirements.

have regard to the user's right to freedom of expression and the right to receive information.

The regulatory landscape for online services is evolving

2.20 The new VSP Framework is part of an evolving and interrelated landscape of online regulations in the UK and internationally.

2.21 Ofcom has been the regulator of video-on-demand services (known as "on-demand programme services" or "ODPS") since 2010. The level of control that a provider exercises over video content available on the service is the key differentiating factor between an ODPS and a VSP. ODPS providers are deemed to have general control because they actively select the content that is available on their services; whereas VSP services provide the functionality for videos to be uploaded by their users. As a result, the responsibilities placed on providers to ensure their users are appropriately protected differ under the respective ODPS and VSP regulatory frameworks.

2.22 It is possible for platforms to offer both a distinguishable ODPS and VSP service, or to be predominantly a VSP service which carries an ODPS. Service providers should consult published Ofcom ODPS Guidance and VSP Scope Guidance to understand more.

2.23 In December 2020, the Government confirmed its intention to appoint Ofcom as the regulator of the future online harms regime and re-stated its intention for the VSP Framework to be superseded by new legislation following the commencement of the online harms regulatory framework.[11]

2.24 Ofcom will operate the VSP Framework until such time as it is no longer in force and will ensure that there is support for services transitioning. Ofcom views the VSP regime as an important precursor to the future Online Safety legislation. Given the two regimes' shared objective to improve user safety by requiring services to protect users through the adoption of appropriate systems and processes, Ofcom considers that compliance with the VSP regime will assist services in preparing for compliance with the online harms regime as described by Government in its response to the Online Harms White Paper.

2.25 VSP regulation will also sit closely alongside other regulatory regimes such as data protection and the Information Commissioner's Office (ICO)'s Age Appropriate Design Code (AADC). The VSP Regulations seek to protect users, particularly under-18s, from harmful material, while the AADC sets out how data protection by design and data protection law apply in respect of online services likely to be accessed by children. These are often closely connected issues. Ofcom is working closely with the ICO, as well as the Competition and Markets Authority (CMA), through the Digital Regulators Cooperation Forum to support regulatory coordination in online services and cooperate on areas of mutual importance.[12]

[11] Full government response to the Online Harms White Paper
[12] DRCF Launch Document

The requirements of the VSP Framework

2.26 The VSP Framework sets out a number of requirements on VSP providers. These include administrative requirements such as:

a) **Notifying Ofcom.** The obligation to notify will come into force on 6 April 2021 and existing UK-established VSP providers will have until 6 May 2021 to notify their service to Ofcom. Services commencing after 6 April are required to make an advance notification to Ofcom of their intention to provide a service. See our statement on who needs to notify for more information.

b) **Publishing information.** This includes the VSP provider's name, address and email address; a statement that the VSP provider is under the jurisdiction of the UK and; Ofcom's name, address and email address.[13]

c) **Paying Ofcom a fee**. Ofcom may set and charge VSP providers an annual fee from April 2022. We will consult providers before setting any fee.[14]

d) **Complying with information requests and co-operating with Ofcom.**[15]

Users should be protected from harmful material

2.27 Two of the principal objectives of the VSP Framework are to protect the general public from relevant harmful material and to protect under-18s from restricted material. Here we briefly set out the statutory definitions for these two categories of harmful material and provide further information in Section 3.

2.28 VSP providers must take such of the measures listed in the legislation as are appropriate for the following purposes:

a) Protecting persons under the age of 18 from videos and adverts containing **restricted material**.

> **Restricted material** refers to videos which have or would be likely to be given an R18 certificate, or which have been or would likely be refused a certificate.[16] It also includes other material that might impair the physical, mental or moral development of under-18s.

b) Protecting the general public from videos and adverts containing **relevant harmful material.**

[13] Section 368Y (2) of the Act
[14] Sections 368Y (3) (a) and Section 368Z9 of the Act
[15] Sections 368Y (3) (b) and (c) of the Act.
[16] Certificate here refers to 'classification certificate' which has the same meaning as in the Video Recordings Act 1984. See BBFC Classification Guidelines(pages 28 – 31) for more information on R18 certificates and the refusal to classify works.

Relevant harmful material refers to any material likely to incite violence or hatred against a group of persons or a member of a group of persons based on particular grounds.[17]

It also refers to material the inclusion of which would be a criminal offence under laws relating to terrorism; child sexual abuse material;[18] and racism and xenophobia.

2.29 Throughout this Guidance, we use the term "harmful material", where appropriate, to refer to both restricted material and relevant harmful material. We also occasionally refer to the requirements to have in place appropriate measures to protect the general public from relevant harmful material and protect under-18s from restricted material collectively as "the requirement to protect users from harmful material".

Standards around advertising

2.30 The requirement to have in place appropriate measures to protect users from harmful material applies to all videos,[19] whether or not they are, or include, adverts.[20] VSP providers are obliged to comply with these requirements whether or not the adverts have been marketed, sold or arranged by the VSP provider.

2.31 The VSP Framework also includes requirements to ensure adverts on VSPs comply with specific advertising requirements around transparency, prohibited and restricted products and other general advertising requirements. These "advertising-specific requirements" vary depending on whether or not the adverts have been marketed, sold or arranged by the VSP provider.

2.32 Ofcom will consult separately on these advertising-specific requirements, including guidance on the application of "marketed, sold or arranged" and a proposal to work with the Advertising Standards Authority (ASA). As a result, this Guidance does not cover the advertising-specific requirements.

VSP providers should take appropriate measures to protect users from harmful material

2.33 Below we set out the measures listed in the VSP Framework which platforms can take in order to protect users from the harmful material defined above. It is important to note

[17] The particular grounds are: grounds such 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. The VSP Framework currently refers to Article 21 of the Charter of Fundamental Rights of the European Union where these grounds are set out. Under Government regulations which have been laid before Parliament, the reference to Article 21 is due to be replaced with the grounds listed in Article 21.

[18] The AVMSD uses the phrase "the sexual abuse and sexual exploitation of children and child pornography".

[19] "Video" is defined as a set of moving or still images, or of legible text, or of a combination of those things (with or without sounds), which constitutes an individual item irrespective of its length (and which is not an audiovisual commercial communication).

[20] The legislation refers to "audio-visual commercial communications" ("AVCCs"). AVCCs is a term applied across a number of sectors and includes advertising, as well as sponsorship, teleshopping and product placement, but also influencer marketing and other forms of commercial communication associated with VSPs. In this guidance, "adverts" and "advertising" are used as a short-hand for "AVCCs".

that where a measure is taken, it must be implemented in such a way as to carry out the requirement to protect under-18s from restricted material and/or the general public from relevant harmful material.[21]

1. Include **terms and conditions** to the effect that if a person uploads a video that contains any **restricted material**, that person must bring it to the attention of the VSP provider.
2. Include **terms and conditions** to the effect that a person must not upload a video containing **relevant harmful material**.
3. Include **terms and conditions** about the requirements of **adverts** on the platform.
4. Provide the **functionality** for someone uploading a video to declare whether the video contains an **advert**.
5. Establish and operate transparent and user-friendly mechanisms for viewers to **report or flag harmful material** and provide explanations to users about any action taken in response to material that has been reported or flagged by viewers.
6. Establish and operate easy to use systems allowing **viewers to rate harmful material**.
7. Establish and operate systems for obtaining **assurance as to the age of potential viewers**.
8. Provide for **parental control systems** in relation to restricted material.
9. Establish and operate a **complaints procedure** in relation to the implementation of: reporting or flagging mechanisms and in relation to the outcome of any action taken in response; age assurance systems; rating systems; and parental controls in relation to restricted material. This must be transparent, easy to use and effective, and must not affect the ability of a person to bring a claim in civil proceedings.
10. Provide tools and information for users with the aim of improving their **media literacy** and raise awareness of the availability of such tools and information.

2.34 Measures 3 and 4 above relate to advertising on VSPs and are more directly related to the advertising-specific requirements set out at 2.31 above. As such, they are not covered in this document.

2.35 We occasionally refer to the other eight measures collectively as “protection measures” and provide more detail on these in Section 4, including how VSP providers might consider implementing the measures.

2.36 Whether a measure is appropriate for the purposes of protecting users from harmful material, must be determined by whether it is practicable and proportionate (see Section 5). This requires VSP providers to take into account particular factors which include the type of harm that may be caused, the characteristics of those whom the measure is

[21] Section 368Z1 (2) of the Act

intended to protect (e.g. under-18s or any other category of user), the size and nature of the service and the rights and legitimate interests of users.

How to use this guidance

2.37 This guidance is designed to help VSP providers understand what is expected of them under the VSP Framework. It is open to providers to determine which measures they take and how they implement them to protect users from harmful material. In regard to this requirement, the VSP Framework specifies that:

a) Providers should consider the measures set out in the VSP Framework. Section 4 of this Guidance sets out these measures.

b) Where measures are taken, they must be implemented in such a way as to meet the requirements of the regime (i.e. to protect users from harmful material). Section 4 of this Guidance provides some ways in which this might be achieved.

c) Providers determine which of the measures to take according to whether it is practicable and proportionate.

i) Section 5 of this Guidance explains the practicable and proportionate criteria set out in the VSP Framework, including how they might impact a provider's decisions about which measures to take and how to implement them.

ii) Section 6 of this Guidance encourages the use of additional steps to help protect users. These include considering the practicable and proportionate criteria and decisions about protection measures as part of a risk management process. This process should involve identifying potential harms on a platform; documenting decisions about the measures in place to mitigate those potential harms; and measuring the effectiveness of those measures.

2.38 We recognise that there are significant differences between the platforms in scope of the VSP Framework and we understand the importance of innovation in the safety tech sector. As such, this document should not be viewed as a set of compulsory steps. Providers should use this document to support their considerations on which measures are appropriate for their service to secure the required protections.

2.39 Where we think effective protection of users is unlikely to be achieved without a specific approach, we say so. However, providers must ensure they comply with their statutory obligations under the VSP Framework and should seek their own legal advice on any compliance issues arising.

2.40 Where a provider is able to clearly demonstrate that they are taking compliance with their requirements seriously, including following the suggestions in this document, and can evidence users being appropriately protected from harmful material, there is a greater likelihood Ofcom will consider those platforms to be in compliance. Section 7 has further details about Ofcom's approach to monitoring compliance and enforcement.

2.41 The structure of the remainder of the Guidance is as follows:

- **Section 3** sets out guidance concerning the **harmful material** users are to be protected from, as defined in the regulations;
- **Section 4** sets out guidance concerning the **protection measures** in the VSP Framework, including how they might be implemented to effectively protect users from harmful material;
- **Section 5** sets out guidance concerning the **practicable and proportionate criteria** to be considered when VSP providers are deciding which measures to take to protect users on their specific platform;
- **Section 6** lists **additional steps to protect users** which are not explicitly covered by VSP Framework. Included in this section is guidance on risk management processes which we encourage platforms to conduct; and
- Section 7 sets out Ofcom's approach to monitoring and enforcement.

3. Harmful material

3.1 The VSP Framework requires VSP providers to take measures (set out at 2.33 above) to protect users from certain categories of harmful material. This section considers the different types of harmful material that are likely to be caught under the relevant definitions, drawing on research commissioned by Ofcom.

3.2 As noted in Section 2, the VSP Regime focusses on the measures that providers take to protect their users and does not set standards for the content of individual videos. Therefore, Ofcom will not usually review individual pieces of content when investigating matters of compliance, but we may need to have regard to them as part of our supervision or enforcement activities.

3.3 We expect providers to be aware of the definitions of the categories of harmful material in the VSP Framework, to support their implementation of protection measures and any of their own assessments, for example in response to reports of harmful material appearing on the service. Providers can also use this information to support any risk management processes they have in place (see Section 6).

3.4 Harmful Material encompasses **restricted material** and **relevant harmful material.** The definitions of these terms are below.[22]

a) **Restricted material** refers to:

i) Videos which have, or would be likely to be given, an R18 certificate.[23]

ii) Videos containing material not suitable for BBFC classification.[24]

iii) Material that might impair the physical, mental or moral development of under-18s.

b) **Relevant harmful material** refers to:

i) Material likely to incite violence or hatred against a group of persons or a member of a group of persons based on any ground such 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 or sexual orientation.

ii) Material the inclusion of which would be a criminal offence under laws relating to terrorism; child sexual abuse material; and racism and xenophobia.

[22] See Section 368Z1(8) and paragraph 10 in Schedule 15A of the Act.
[23] Certificate here refers to BBFC 'classification certificate'. There is no requirement for material being provided on an VSP to be classified by the BBFC, but Ofcom is required to have regard to the BBFC Classification Guidelines when determining whether material on a VSP is R18-equivalent.
[24] The British Board of Film Classification is responsible for 'classification certificate' which has the same meaning as in the Video Recordings Act 1984.

Restricted material

Videos which have or would be likely to be given an R18 certificate

3.5 The R18 category is a special and legally-restricted classification, primarily for explicit videos of consenting sex or strong fetish material involving adults, and where the primary purpose of the material is sexual arousal or stimulation.

3.6 This includes material which has an R18 classification certificate, as well as material whose nature is such that it is reasonable to expect that if it was submitted to the BBFC for a classification certificate, the BBFC would issue a R18 classification certificate. [25]

Material unsuitable for classification

3.7 Restricted material includes material which has either been determined not suitable for a classification certificate by the BBFC or material whose nature is such that it is reasonable to expect that it would not be suitable for a classification certificate.[26]

3.8 The BBFC's current guidelines outline that material likely to be unsuitable for classification could include: material which is in breach of criminal law (or created through the commission of a criminal offence); material that appears to risk harm to individuals or to society such as, for example, the detailed portrayal of violence or dangerous acts, illegal drug use; and portrayal or invitations to conduct sadistic violence, rape or other non-consensual sexual violent behaviour or other harmful violent activities.[27]

3.9 While the VSP Framework does not require videos on VSPs to obtain a BBFC classification, providers should regularly consult BBFC guidelines to understand the type of material that is likely to be unsuitable for classification.

Other material that might impair the physical, mental or moral development of under-18s

3.10 The type of material online that might impair the physical, mental or moral development of under-18s is vast and likely to evolve as user behaviour and services adapt.

3.11 The legislation does not specify any particular examples of material that might impair the physical, mental or moral development of under-18s. In order to support a greater understanding of this, Ofcom commissioned a wide-ranging research study into the risks and harms to children and young people being online, using social media and VSPs.[28] The report goes beyond the VSP Framework but providers may find it helpful to consider the report's findings. In particular, the following potential harms could be relevant to consider

[25] See BBFC Classification Guidelines on R18 material
[26] Section 368E (3) (a) and (b) of the Act
[27] See BBFC Classification Guidelines.
[28] See UEL Report for more examples and information on the academic evidence to support these categories of harm.

when determining which measures it may be appropriate to take, to protect under-18s from material that might impair the physical, mental or moral development:

a) **Sexual material**, including pornography,[29] sexting, naked selfies and nudes, grooming and meeting strangers online;

b) **Self-injurious content** which may cause physical harms, such as material promoting eating disorders, self-harm and suicide;

c) **Mental health and wellbeing factors** which may lead to a harm, such as psychological distress, depression, anxiety, social withdrawal, body image and addictive-type behaviours;

d) **Aggression**, including hate speech, violent material, dangerous behaviour, cyberbullying, online harassment, and cyberstalking;

e) **Manipulation intended to harm**, through image, AI and algorithmic manipulation; profiling and persuasive design including nudging and targeting leading to a detrimental impact on under-18s.

3.12 For material that might impair the physical, mental or moral development of under-18s, the principle applies that material that has the most potential to harm must be subject to the strictest access control measures.

3.13 In considering material that might impair the physical, mental and moral development of under-18s, VSPs are advised to take account of the different cognitive abilities of children and young people under the age of 18, who are more vulnerable than adults and may require distinct protections by age range.

3.14 VSPs should also consider whether the material is age-appropriate for its users. To support this approach, it may be useful to understand the strength and types of material that the BBFC regards as appropriate for different age groups in its classification guidelines.[30]

3.15 VSP services created for infants and young children should endeavour to achieve the highest degree of safety, including applying the most robust protection mechanisms and supervision features to involve the parent or carer.

3.16 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.

3.17 Services should ensure they understand the wider online advertising requirements that apply under the rules enforced on a self-regulatory basis by the Advertising Standards Authority (ASA) and data protection rules that apply to under-18s, enforced by the ICO. [31]

[29] Videos whose primary purpose is sexual arousal or stimulation should be considered as only suitable for adults, in line with BBFC Classification Guidelines on Sex works at 18.
[30] See age ratings issues in BBFC Classification Guidelines
[31] The Cap Code: The UK Code of Non-Broadcast Advertising and Direct & Promotional Marketing

Services likely to be accessed by children should ensure they meet the data protection requirements in the ICO's Age Appropriate Design Code.[32]

3.18 Overall, VSPs should aim to balance protecting under-18s from material that might impair their physical, mental and moral development with the benefit to children and young people when using the service to acquire knowledge, connect with others, and seek enjoyment and self-expression.[33]

Relevant harmful material

3.19 The general public (i.e. all VSP users) must be protected from relevant harmful material. Some of the material included under that definition relates to criminal offences and services should seek legal advice on applicable laws for such material.

Material likely to incite violence or hatred

3.20 The VSP Framework includes, as part of the definition of relevant harmful material, material likely to incite violence or hatred against a group of persons, or a member of a group of persons, based on any ground such 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 or sexual orientation.[34]

3.21 Services need to be aware that whether content is likely to incite violence or hatred will vary depending on the nature of the protected characteristic, the stereotypes that exist and the social context. Where it relates to a protected group, publicly condoning, denying or grossly trivializing crimes of genocide, crimes against humanity, war crimes and crimes against peace (the planning or carrying out of a war of aggression) may also amount to incitement to violence or hatred.[35]

3.22 "Incitement to hatred" should be understood as having its usual meaning in everyday language.[36] When determining the appropriateness of the measures VSPs take, Ofcom will have regard to relevant European case law on freedom of expression, which includes the case law of the European Court of Human Rights (ECHR).[37] In September 2020 the ECHR published a factsheet summarising some of its cases on incitement to hatred, which may be helpful to providers.

3.23 Separately, Ofcom commissioned The Alan Turing Institute to produce a report on online hateful content. The report goes beyond the VSP Framework, looking at the nature, dynamics and prevalence of this harm on the internet. The examples in the report of

[32] For further information, see the ICO Children's Code Hub
[33] See UEL Report for further information on the benefits of the Internet for children and young people.
[34] The VSP Framework currently refers to Article 21 of the Charter of Fundamental Rights of the European Union where these grounds are set out. Under Government regulations which have been laid before Parliament, the reference to Article 21 is due to be replaced with the grounds listed in Article 21.
[35] European Council Framework Decision 2008/913/JHA
[36] *Mesopotamia Broadcast A/S METV v Germany* C244/10 and C245/10
[37] ECHR cases can be found here

industry approaches to addressing online hate may be relevant for services considering their own approach to hateful content.[38]

Material the inclusion of which would be a criminal offence

3.24 The VSP Framework requires the general public (i.e. all VSP users) to be protected from material which would be a criminal offence to publish, distribute or disseminate under laws relating to terrorism; child sexual abuse material (CSAM); and racism and xenophobia.

3.25 Where this type of material is involved, it is good practice for VSPs to establish internal protocols to quickly identify, escalate and action this content if it is required. To ensure such systems are effective and appropriate to the type of content, VSPs may wish to seek the advice of relevant authorities, including from UK law enforcement agencies on how to deal with illegal content (e.g. the National Crime Agency).[39]

3.26 VSPs should also consider collaborating or partnering with organisations that work with online platforms on terrorism and child sexual abuse material – for example, with the Internet Watch Foundation (IWF), the National Centre for Missing and Exploited Children (NCMEC), or Thorn for CSAM; and with Tech Against Terrorism or the Global Internet Forum to Counter Terrorism (GIFCT) for terrorist material. Where practical, we encourage providers to explore more formal partnerships with these organisations, particularly for platforms with higher risk profiles for this type of content.

Terrorism

3.27 The VSP Framework refers to Section 1 of the Terrorism Act 2006, 'Encouragement of Terrorism'.[40] A person commits an offence under this section if they publish a statement directly or indirectly encouraging terrorism or that is likely to be understood as such, whether intentionally or recklessly. It is irrelevant whether any person is in fact encouraged or induced by the statement to commit, prepare or instigate a terrorist act.

3.28 Statements that are likely to be understood as indirectly encouraging the commissioning or preparation of acts of terrorism include every statement which glorifies the commission or preparation (whether in the past, in the future or generally) of such acts or offences.

Child Sexual Abuse Material

3.29 The VSP Framework refers to Article 5(4) of Directive 2011/93/EU of the European Parliament and of the Council of 13 December 2011 on combating the sexual abuse and sexual exploitation of children and child pornography ("the CSEA Directive").

[38] See the Turing Report.
[39] For the avoidance of doubt, platforms should not send potentially illegal content to Ofcom.
[40] The definition of relevant harmful material in the VSP Framework currently refers to Article 5 of Directive (EU) 2017/541 of the European Parliament and of the Council of 15 March 2017 on combating terrorism. Under Government regulations which have been laid before Parliament, this reference to EU legislation is due to be replaced with a reference to section 1 of the Terrorism Act 2006.

3.30 We consider the offences under the CSEA Directive most relevant for VSP providers to be related to the distribution, dissemination or transmission of child pornography.[41] The definition of child pornography is set out in Article 2 of the CSEA Directive and extends to the depiction of any person appearing to be a child as well as realistic images. It also includes simulated activity.

3.31 In relation to this type of content, VSPs may wish to familiarise themselves with the Memorandum of Understanding agreed between the Crown Prosecution Service (CPS) and the Association of Chief Police Officers (ACPO), which may be helpful when planning what to do if CSEA content is flagged on their platforms.

Racism and Xenophobia

3.32 The VSP Framework refers to Article 1 of Council Framework Decision (2008/913/JHA) of 28 November 2008 on combating certain forms and expressions of racism and xenophobia by means of criminal law.

3.33 The offences relating to racism and xenophobia here include publicly inciting violence or hatred directed against a group of persons or a member of such a group defined by reference to race, colour, religion, descent or national or ethnic origin, and the committing of such an offence by public dissemination of distribution of tracts, pictures or other material.

3.34 Also included are offences related to publicly condoning, denying or grossly trivializing crimes of genocide, crimes against humanity, war crimes and other specified crimes,[42] directed against a group or group of persons defined by the characteristics in 3.33 above, where the conduct is carried out in a manner likely to incite to violence or hatred against such a group or a member of such a group.

Consultation question

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

[41] Throughout the rest of the document we refer to these offences under the term “child sexual abuse material” or “CSAM”.

[42] These other specified crimes are crimes defined in Article 6 of the Charter of the International Military Tribunal appended to the London Agreement of 8 August 1945.

4. Protection measures

4.1 VSP providers should take the measures listed in the VSP Framework which they determine to be appropriate for protecting users from harmful material. In this section we provide guidance on the measures in the VSP Framework (see paragraph 2.33) including how they might be implemented in a way that protects users. Guidance on measures that are most relevant to ensuring advertising standards are met will be published separately.[43]

4.2 It is for individual providers to determine whether a measure is appropriate for them to take, having regard to particular factors which include the size and nature of their service, the profile of their users and the nature of the material available. Section 5 provides more detail on this.

4.3 This section first provides guidance on how providers can implement protection measures to meet the requirements of the VSP Framework, by reference to some key principles. It then covers the following:

- Terms and conditions
- Reporting and flagging mechanisms
- Systems for viewers to rate harmful material
- Tagging or rating restricted material
- Age assurance systems
- Parental controls systems
- Complaints process
- Dispute resolution procedure
- Media literacy tools and information

The implementation of protection measures

4.4 Where a VSP provider decides to take one or more of the measures listed in the VSP Framework, those measures must be implemented in such a way as to carry out the relevant purpose.[44] In other words, providers need to ensure that they are implemented in a way that protects under-18s from restricted material and the general public from relevant harmful material.

4.5 The requirement to implement measures in a way that protects users, does not mean that Ofcom expects all harmful material to be eradicated from a platform as a result of these measures. However, one of the aims of effective implementation should be to prevent users from encountering harmful material, potentially by reducing the prevalence of it across a platform, or restricting access to it. Evidence of, for example, continued

[43] The measures that are most relevant to ensuring advertising standards are met are to include terms and conditions to the effect that of advertising-specific requirements are met and to provide the functionality for someone uploading a video to declare whether the video contains an advert.
[44] Section 368Z1 (2) of the Act

occurrence of harmful material appearing on a platform may suggest that a platform has not taken appropriate measures or has not implemented them effectively.

4.6 In this section we have provided guidance on what platforms "should do" or "should consider" when implementing measures in a way that achieves the requirement to protect users. These are not prescriptive requirements but intended as helpful suggestions to aid understanding of how compliance could be achieved. In some instances, there may be other ways to implement a measure to achieve the same requirement.[45] Where we think effective protection of users is unlikely to be achieved without a specific approach, we say so.

4.7 The guidance on the implementation of these measures is intended to apply to as many services as might fall under the definition of a VSP. Individual characteristics of different services might affect considerations about the implementation of measures and providers should have regard to the practicable and proportionate criteria in Section 5 alongside the guidance in this section.

Five principles to support implementation

4.8 The VSP Framework sets out specific requirements as to how certain measures need to be established and operated. For example, reporting and flagging mechanisms must be **transparent** and **user-friendly** and systems allowing viewers to rate harmful material must be designed so that they are **easy to use**. Additionally, complaints processes must be **transparent, easy to use** and **effective**.

4.9 These requirements are not mandated in relation to other measures, such as terms and conditions, however, we consider it good practice for providers to take these principles into account in designing and implementing all of their protection measures. We also suggest it may be helpful to consider **fairness** and the need for measures to **evolve** or adapt. We provide further detail on these principles below:

[45] We recognise that there may also be other measures not listed in the VSP Framework which may achieve the same protections.

Effective: measures should be implemented in a way that achieves the objective of protecting users. This includes taking necessary steps to operate and apply those measures (e.g. terms and conditions cannot be effective if they are not enforced).

Easy to use: measures employed by platforms should be easy for all users to locate, understand and engage with.

Transparent: the intended and actual outcomes of any measures taken by a platform should be clear to users and other interested stakeholders.

Fair: measures should be designed and implemented in a way that does not unduly discriminate between users, introduce bias or result in inconsistent application.

Evolving: it is good practice to ensure that measures are regularly reviewed and updated in line with changing user behaviours and technological advancements to ensure that they remain appropriate for their intended purpose.

4.10 We suggest that these five principles are overarching in their application, although where relevant we have drawn particular attention to individual factors when they are most relevant to a particular measure.

Measuring effectiveness

4.11 Platforms wishing to assure themselves that they are in compliance with the VSP Regime should assess the effectiveness of their measures. We suggest how effectiveness might be assessed for different protection measures in this section. Such information collection is not a requirement, but is likely to support Ofcom's understanding of how measures are implemented and whether they are appropriate for protecting users. More details on the information Ofcom might request to assess effectiveness can be found in Section 7.

Ease of use for vulnerable users

4.12 Some of the measures are required to be easy to use. Even where this is not the case, we encourage providers to consider the needs of vulnerable users when designing or implementing particular measures. Vulnerability might be related to physical or mental health problems or specific personal circumstances such as age or literacy skills. The following are among the various techniques that could be considered: using simple language, not overcrowding information, highlighting or emboldening key pieces of information and making sure information is accessible, for example that it is readable by screen reader software.

Fairness and user's rights

4.13 Providers must have regard to the rights and legitimate interests of the users of their service. Here we encourage providers to consider how they balance the effectiveness of the protection measures described below, with the rights of users. This principle is relevant both for the implementation of measures (which is covered in this section) and the

decisions about which measures to take (considerations about which are covered in Section 5).

4.14 We encourage providers to design and implement measures in a way that does not unduly discriminate between users, introduce bias or result in inconsistent application. We are particularly mindful of the risk of over-takedown or unnecessary censorship on platforms when, for example, enforcing terms and conditions by removing content in response to flagging or reporting. Complaints processes and dispute resolution procedures will play an important role here (see 4.114 – 4.146).

Terms and conditions

4.15 The first two measures listed in the VSP Framework relate to the inclusion of terms and conditions for users governing the uploading of material. These require that:

- videos containing restricted material are brought to the attention of the provider of the service; and
- videos containing relevant harmful material must not be uploaded to the service.

4.16 Terms and conditions may include community guidelines; community standards; terms of service; or any other rules or standards used to govern the type of content permitted on a VSP.

4.17 Below we provide guidance on terms and conditions about harmful material and how the corresponding measures may be implemented to ensure they are appropriate for protecting users.

Terms and conditions about restricted material

> *Include terms and conditions to the effect that if a person uploads to the service a video that contains any **restricted material**, that person must bring it to the attention of the person who is providing the service.*

4.18 Terms and conditions should explain the type of content considered to be restricted material (with reference to Section 3 of this Guidance) and specify that videos containing this material must be brought to the attention of the VSP provider.

4.19 It is for VSP providers to determine the method by which a person uploading a video can bring it to the attention of the provider, including whether binary tags or more graded rating systems are appropriate. Providers should also consider what action they take to protect under-18s in response to being made aware of videos containing restricted material.

4.20 We are of the view that, where providers choose to take this measure, it is unlikely that effective protection of under-18s can be achieved without the provider taking the additional step of either notifying viewers where a video contains restricted material or restricting access to it by under-18s.

4.21 This could potentially include employing tagging and rating mechanisms (see 4.79 – 4.86) or employing access control measures. Access control measures refer to any measure which restricts a user's access to content. Age assurance and parental controls are the two access control measures listed in the VSP Framework and are described below from 4.89.

4.22 We recognise that for some types of platform, the inclusion of terms and conditions requiring users to notify the provider if they upload restricted material may not be necessary. For example, if the platform specialises in restricted material of a pornographic nature. In such cases we expect providers to consider taking and implementing other measures to protect under-18s from videos containing restricted material, such as having appropriately robust age assurance systems in place (see 4.87 – 4.108).

Terms and conditions about relevant harmful material

> *Include terms and conditions to the effect that a person must not upload to the service a video containing* ***relevant harmful material.***

4.23 As noted above, relevant harmful material refers to any material likely to incite violence or hatred against a group of persons or a member of a group of persons based on particular grounds.[46] It also refers to material the inclusion of which would be a criminal offence under laws relating to terrorism; child sexual abuse material; and racism and xenophobia.

4.24 Types of content which fall under the definition of relevant harmful material are set out above in Section 3.

4.25 Ofcom considers this measure as fundamental to the VSP Regime and we consider it is unlikely that effective protection of users can be achieved without having this measure in place and it being implemented effectively.

Ensuring terms and conditions are easy to use

4.26 The easier it is for users to be able to locate, understand and engage with a provider's terms and conditions, the more effective they are likely to be in helping to protect users from harmful material. In considering this, VSP providers should have regard to the length; readability; location; format; timing and promotion of terms and conditions.

Length and readability

4.27 Long and complex terms and conditions mean users are unlikely to engage with them and this is unlikely to lead to their effective implementation, particularly on platforms which are popular with under-18s.

4.28 Therefore, we consider it is important that platforms ensure that all terms and conditions can be easily understood by users, for example by avoiding the use of jargon and legalese

[46] The particular grounds are: grounds such 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.

and providing clear explanations of key concepts and requirements. This is particularly so given that around 15% of UK adults have poor literacy skills.[47]

4.29 Further, platforms who have a typically younger user profile should consider providing child-friendly explanations. We note that studies involving such simplified terms and conditions have demonstrated increased engagement from younger users.[48]

4.30 Platforms should also consider the needs of vulnerable users when drafting and reviewing terms and conditions. Charities and other organisations can provide resources to help consider accessibility requirements.[49]

Location, format, timing and promotion

4.31 Platforms should consider how and when terms and conditions are accessed by users. Typically, terms and conditions are presented to a user upon signing up to a service and can be found in 'Help' or Settings' areas of a website or application. It should be easy for any user to locate the terms and conditions.

4.32 We also consider that the method of accessing terms and conditions is important to user engagement. For example, it is unlikely to be appropriate for an app-based VSP to direct users to pages on a website designed to be accessed primarily on a desktop computer.

4.33 We note that on some platforms, users are required to review the terms and conditions when an account is created but are not prompted to engage with them again subsequently. This is unlikely to ensure users remain up-to-date with policy changes and that they are aware of changes to guidelines about what type of content can be uploaded.

4.34 VSP providers should consider how frequently users should be prompted to engage with terms and conditions based on a platform's own risk profile (see Section 6). Prompts could be used, for example, to remind users of relevant parts of the terms and conditions at different points in the user journey, such as when uploading a video.

4.35 The Behavioural Insights Team (BIT) produced an evidence-based best practice guide on presenting terms and conditions to consumers. The aim of this was to improve the number of people who open terms and conditions, read and understand them.

4.36 Some of the factors that were found to lead to an increase in the number of people who open terms and conditions include: showing users the full text of the terms and conditions within a scrollable text box (instead of requiring a click to view them); telling users how long it will take to read the terms and conditions; and telling users when it is the last opportunity for them to review terms and conditions before making a decision (e.g. signing up). The same study also found that summary bullet-points with icons illustrating key terms can lead to increased understanding of terms and conditions. [50]

---

[47] See the National Literacy Trust's webpage on adult literacy levels
[48] "Growing Up Digital", Children's Commissioner, January 2017
[49] For example, CHANGE's 'How To Make Information Accessible' guide
[50] Improving consumer understanding of contractual terms and privacy policies: evidence-based actions for businesses, The Behavioural Insights Team.

Ensuring terms and conditions are effective

4.37 Terms and conditions need to be implemented in such a way to meet the requirement of protecting users. In practice, this is likely to mean having robust processes in place to ensure that terms and conditions are appropriately enforced. A clear way of achieving this would be through content moderation and appropriate sanctions for violations.

Enforcement and sanctions

4.38 To work effectively as a protection measure, terms and conditions need to be appropriately enforced through a process owned or overseen by the VSP provider. Violations of the terms and conditions should result in effective action being taken by the VSP.

4.39 Individual providers will need to decide what action may be appropriate for particular circumstances and the thresholds for taking action. Serious violations and repeat offences should attract the toughest sanctions.

4.40 Effective action in response to violations might include warnings; temporary bans on posting content; bans on interacting with the content of others; demonetisation; temporary account restrictions; and permanent removal or deletion of accounts. We are aware that some VSPs also block IP addresses.

Moderation

4.41 VSP providers can choose different ways of enforcing terms and conditions on their platforms. Many involve some form of moderation, either by humans, or by employing machine-learning, or both. This can occur prior to, during, or very recently after the uploading process (proactive moderation), or in response to reports from users (reactive moderation).

4.42 Ofcom expects any moderation techniques used to be subject to regular quality assurance processes, so that they remain effective at enforcing the terms and conditions of a platform and, ultimately, provide effective protection to users from harmful material. Ofcom would expect the accuracy of any machine-learning moderation techniques to be checked using human quality assurance processes.

4.43 The size and operation of teams involved in the moderation of content on a VSP will vary dependent on the size and nature of the platform itself (see Section 5). Some VSPs have safety teams operating 24/7. We consider it best practice for providers, when considering moderation, to reflect the global reach of their service and the constant engagement of their users in accessing and uploading content.

4.44 Where VSP providers use external moderation services to assist the identification and removal of harmful content, they should ensure they are effective. Ultimately, the responsibility for ensuring that protection measures are effective at protecting users from harm lies with the VSP provider, even where such third parties are involved.

4.45 We encourage VSPs to collect information to assess the effectiveness of any moderation techniques and to support any risk assessments (see Section 6). Ofcom may use this information to support its compliance monitoring (see Section 7).

Ensuring the implementation of terms and conditions is fair and transparent

4.46 We recognise that it is often essential for moderation techniques to remain confidential, to better allow VSP providers to effectively remove harmful material. However, in order to be transparent, the consequences of breaching terms and conditions (including the sanctions mentioned above) need to be clear to all users.

4.47 Taking into account the rights and legitimate interests at stake, providers should ensure that terms and conditions are enforced in a manner that does not unduly discriminate between users, introduce bias or result in inconsistent application.

4.48 Where actions such as content removal, blocks or bans occur, users need to have the ability to understand and challenge these decisions. This is vital to ensure that users (including users which are regulated broadcasters or other media outlets) are not unreasonably impacted by over-takedown of content. Careful consideration should also be given to claims which involve videos containing news content and we explain this further in paragraph 4.145 below.

4.49 The complaints process and dispute resolution procedure set out below are the two primary functions that will help VSP providers ensure moderation and enforcement decisions are fair (see 4.114 – 4.146). Terms and conditions about harmful material should clearly signpost these functions.

Ensuring terms and conditions continually evolve

4.50 In order for terms and conditions to continue to protect under-18s from restricted material and protect the general public from relevant harmful material, we expect VSP providers to keep terms and conditions under review and make changes where necessary.

4.51 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.

4.52 Where appropriate, providers should ensure there is sufficient opportunity for users to understand the nature and impact of any change before it is implemented. Training or guidance for creators on the updated terms and conditions might be considered in order to improve compliance with them.

4.53 VSP providers should also be aware of potential changes to the ways in which users interact with their platform as these might change the risk profile of the platform and terms and conditions (as well as policies surrounding their implementation) may need to adapt to reflect this.

Consultation questions

Question 2: Do you have any comments on the draft guidance about measures which relate to terms and conditions, including how they can be implemented?

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?

Question 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 be taken in response to this notification to achieve effective protection of under-18s, such as applying a rating or restricting access?

Reporting and flagging mechanisms

Establish and operate:

(a) transparent and user-friendly mechanisms for viewers to report or flag harmful material which is available on the service to the person providing the service; and

(b) systems through which the person providing the service explains to the persons using the service what effect has been given to the reporting and flagging referred to in sub-paragraph (a)

4.54 Reporting and flagging mechanisms enable users to notify a VSP provider about harmful material and any content that may violate the platform's terms and conditions. In the context of VSP regulation, harmful material refers to video content only. We note that some platforms also allow users to report and flag non-video content (e.g. comments and direct messages), however this type of communication is not generally expected to be caught by the statutory definition of "video".[51]

4.55 Reporting and flagging mechanisms may include: three dots, flag, or similar icons situated near a video; reporting functions embedded in a video; and allowing viewers to contact platforms directly about content concerns, for instance by email or via an online form.

4.56 Reporting and flagging mechanisms can be used by platforms for the purposes of protecting users from harmful material, but may also be used by platforms to allow users to report adverts which have not met the advertising-specific requirements, which Ofcom will consult on separately.

4.57 The two measures in the grey box above are independent of one another. So where a provider takes the measure to establish and operate reporting or flagging mechanisms, it is

[51] "Video" is defined as a set of moving or still images, or of legible text, or of a combination of those things (with or without sounds), which constitutes an individual item irrespective of its length (and which is not an audiovisual commercial communication).

not then a requirement to establish and operate a system to explain the effect of that report, though we note this second measure is likely to be good practice.

4.58 Ofcom considers reporting and flagging mechanisms fundamental to the protection of users and we consider it is unlikely that effective protection of users can be achieved without having this measure in place and it being implemented effectively.

Reporting and flagging mechanisms should be easy to use

4.59 These mechanisms should be prominent and easy for users to find. If reporting and flagging tools are not immediately visible (e.g. embedded in a video), we would expect platforms to take steps to make users aware of this feature and where it can be found.

4.60 VSP providers with a high number of under-18s using their service should consider this when designing or reviewing their reporting/flagging systems.

Actions taken in response to reports or flags should be clear and transparent

4.61 The likely actions that a platform will take in response to a flag should be apparent to all users, even those who have not yet engaged with the functionality. In practice this means the process should be explained somewhere easily accessible on the VSP.

4.62 Where a provider gives explanations about the effect of flagging and reporting, we encourage providers to actively inform users about the process as they engage with it, including the actions that may be taken as a result of the flag or report.

4.63 Responses from the VSP provider to reporting and flagging may include: moderators reviewing reported/flagged content; removing or reclassifying reported/flagged content; and sanctioning users whose content has been reported/flagged.

4.64 Ofcom research found that although general awareness of safety measures on VSPs is low, flagging and reporting tools are the most widely known, with 60% of VSP users claiming to be aware of this measure. However, 35% of those exposed to potentially harmful material did not take any action. The reasons for this included a perception that it would not make a difference.[52] It is therefore important that platforms are clear with users about the outcomes of reporting and flagging and that the overall process is sufficiently simple to encourage user engagement with the measure.

Responses to reports or flags should be appropriately timely

4.65 Responses to flags should be appropriately timely, proportionate to the size, nature and risk profile of the platform (see Section 5). We encourage platforms to set internal timeframes for responses to reports and flags, and to review performance against these regularly. Such information might support Ofcom in assessing the implementation of this measure (see Section 7).

[52] Ofcom, Safety measures on video-sharing platforms survey (quantitative research) 2021.

4.66 Ofcom expects platforms to take the swiftest action in response to reports about the most harmful categories of content. For example, by having expedited processes for the handling of reports about terrorist material. This is particularly important in periods of heightened terror threats.

4.67 Platforms should consider whether it is appropriate to give users an indication of timeframes for responses to flags or reports, both regarding the timeframe for any potential action and the timeframe for providing the user with information about the effect of their report.

4.68 Platforms should have the systems to provide users with information about what the effect of their report has been. Examples of how this can be done include: emailing users to notify them of the outcome of their report/flag; sending users a message on the platform; and providing users with a dashboard or other interface where they can view the current status of content they have reported or flagged.

4.69 When providing a user with a response to a report or flag, VSP providers should make it clear which video their response is referring to, as users may flag or report multiple videos. For example, by referring to the title of the video or a unique reference number for the report/flag.

Reports and flags should be categorised and recorded

4.70 Reporting and flagging mechanisms should enable users to categorise their report by reference to different types of harmful material or by identifying other platform violations (such as underage user accounts). It is also often useful for users to be able to add additional information to support the report. VSPs should seek to find the right balance of creating a streamlined reporting tool that allows users to quickly and effectively report a piece of harmful content, whilst also providing opportunities for users to report content that may need more explanation as to why it is considered harmful.

4.71 These mechanisms might be supported by internal processes such as escalation. Providers may also wish to consider processes for referral of specific categories to specialist organisations or law enforcement where this is appropriate for effective user protection.

4.72 It is advisable to capture relevant data and information to be able to check if these mechanisms are working to protect users. This would include maintaining a record of the type of reported content, how it is handled internally, and keeping the effectiveness of these mechanisms under continuous review. To increase transparency, VSPs could aim to make their decisions public about how and why content has been removed, reinstated or retained.

Consultation question

Question 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?

Systems for viewers to rate harmful material

Establish and operate easy to use systems allowing viewers to rate harmful material.

4.73 This measure is primarily about allowing viewers on a VSP to apply ratings to restricted material. This can assist providers in taking steps to ensure restricted material is appropriately labelled and that under-18s are effectively protected from viewing it.

4.74 In order to implement the restricted material terms and conditions measure effectively, action should be taken where a provider is notified of videos containing restricted material (see 4.20). One of the ways of achieving this is to have a tagging or rating system in place to alert viewers to the inclusion of this type of material. Such a system may be binary or tiered and we explain these systems in more detail below.

4.75 Where a viewer considers a rating to be incorrectly applied, we would expect them to be able to notify the platform of this, so steps can be taken to review the rating and change it as appropriate.

4.76 Platforms might allow viewers to challenge and ultimately change ratings. For example, if enough viewers believe that a different rating should be applied, this might be amended without intervention from the platform being required. Platforms may also use these rating systems to test or improve the algorithms or other mechanisms which recommend content to users.

4.77 "Crowd-sourcing" of the rating of content in this way is not yet widespread amongst VSPs, but we are aware that it has been applied in limited trials involving both uploaders and viewers.[53] There can be risks of accuracy and gaming involved with user-generated rating systems, which may be why few platforms have adopted them to-date. We would therefore caution against using this in isolation of other types of rating system.

4.78 Some VSPs have functionality for users to like or dislike (or upvote and downvote) content and users may use the like/dislike ratio to inform their decisions about watching material which they are unsure of. Such systems might therefore aid the protection of under-18s from restricted material, but we would not expect them to provide adequate protection when used in isolation from other measures.

Tagging or rating restricted material

4.79 As noted above, having terms and conditions requiring that users notify the platform of videos containing restricted material, is only likely to be effective at protecting under-18s from that material if there is a mechanism to make potential viewers aware of this or restrict access to it by under-18s. Providers should therefore consider what action they take to protect under-18s once they have been notified of such material. One way that can be achieved is through tagging or rating systems.

[53] https://www.yourateit.eu/

Different types of tagging or rating systems may be appropriate

4.80 The most basic way in which platforms can use rating systems to aid the protection of under-18s is to have a binary tagging system, where content notified to the platform as restricted material is tagged automatically as, for example, “Mature”.

4.81 Some platforms might consider having more sophisticated tagging or rating systems, where content is labelled with age-appropriate ratings. This could be determined by the user at the point of upload or by the platform using AI or machine-learning.

4.82 A VSP may rely on an existing age ratings framework such as the BBFC ratings system[54], or the PEGI ratings system,[55] administered by the Video Standards Council Rating Board, for interactive game content.

4.83 We expect providers who choose to use existing, established age ratings frameworks on their platforms to also ensure that this is done with the knowledge of the relevant ratings body. This is to promote consistency of established ratings standards, as well as to protect users who will rely on the accuracy of ratings information provided to them on the VSP.

Tagging and rating systems work well alongside other protection measures

4.84 Where platforms use access control measures such as age assurance or parental controls (see below), these are most effective when tied to any tagging or rating system so that users under-18 cannot access restricted material (or parental controls allow the responsible adult to restrict access to this material). Parental controls can also be used in tandem with more sophisticated rating systems involving multiple tiers, to allow an even more appropriate experience on a platform.

4.85 For material which has the most potential to harm under-18s we would not expect a rating system on its own to be a sufficient measure and in our view this will need to be linked to access control measures.

Ensuring tagging and rating systems are transparent and easy to use

4.86 As noted above, when VSPs require users to notify them that a video contains restricted material, they should make it clear to users what constitutes restricted material (with providers having regard to our guidance in Section 3). Where a tiered system is employed by the platform, it is even more important that explanations about which tags or ratings apply to different types of content are made explicit. Clear examples should be considered and providers should have regard to the suggestions about length, readability, location, format and timing at 4.27 – 4.36.

[54] BBFC Classification Guidelines
[55] VSC PEGI Ratings

Consultation question

Question 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?

Age assurance systems

Establish and operate systems for obtaining assurance as to the age of potential viewers.

4.87 **Age assurance** is a broad term that refers to the spectrum of methods that can be used to be informed about a user's age online.[56] The term may also be used to refer to the level of confidence that a platform has in the age of its users. Examples of age assurance cover a range of potential methods, from users self-declaring their date of birth to the use of face-recognition biometrics and computational methods. Other forms of age assurance may include trusted sources that point to a child's age, such as parental verification tools.[57]

4.88 **Age verification** is a form of age assurance where a user's age is established to the greatest degree of certainty practically achievable and is currently therefore considered the strictest form of access control. It is likely to rely on data sources that can secure a high level of confidence in the information provided. Traditional examples of age verification include solutions such as matching a user to their official age on their passport, driving licence or credit card but techniques are also available which do not require the direct exchange of personal identification with platforms.

4.89 There are a range of methods that may be used for age verification and age assurance. These include, but are not limited to:

a) Biometric analysis, such as analysis of facial features, fingerprints and retinal patterns to estimate age;

b) Behavioural analysis, i.e. behaviour patterns of the user on the platform and their interaction with it (e.g. time, location of web use) to determine likely age;

c) Linguistic analysis, i.e. analysis of written language structure to evaluate age;

d) Profiling, such as using a user's past online activity or browsing history to evaluate certain aspects relating to the user;

e) Third-party attribution, such as digital identity solutions, use of data held by third party organisations (e.g. credit card companies) to validate the claimed age of an individual, or single sign-on schemes;

f) Parental control software and mechanisms;

[56] The development of the concept and definition of age assurance has been supported by the government-led Verification of Children Online research project (VoCO). More information on age assurance can be found in the VoCO Phase 2 report (November 2020)

[57] For more detail on how parental verification tools may support better age assurance, refer to Parental control systems below.

g) Password–protected content;

h) Hard identifiers (e.g. passport scans, credit card details).

4.90 In determining an approach to obtaining appropriate assurances as to the age of potential viewers, we encourage VSP providers to conduct a risk assessment of their platform, having regard to the practicable and proportionate criteria (see Sections 5 and 6). This assessment should give particular consideration to the risk of harm posed to under-18s by the type of restricted material on the platform and the prevalence of such material.

4.91 When considering taking any of the protection measures under the VSP Framework providers should have regard to privacy issues and GDPR requirements. This is likely to be of greater consideration for age assurance and age verification measures. We encourage providers to consult the ICO's guidance on UK GDPR requirements[58] and The Age Appropriate Design Code.[59]

Preventing access to restricted material of a pornographic nature for under-18s

4.92 VSP providers are required under the VSP Framework to apply the principle that restricted material that has the most potential to harm the physical, mental or moral development of under-18s must be subject to the strictest access control measures.

4.93 Access control measures are designed to control the ability of individuals to access videos included in a VSP service and the manner of access. Depending on the level of risk posed to under 18s, VSPs should use age verification measures that either operate as an age-gate to block users from the entire platform or to filter material in a way that can protect under-18s.

4.94 Ofcom interprets the statutory principle in 4.92 to mean that **if a VSP has restricted material on its service that is of a pornographic nature, providers should have a robust access control system that verifies age and prevents under-18s from accessing such material**. This applies to VSPs which specialise in pornographic material, as well as services which have a high prevalence of such material.

4.95 Ofcom regards restricted material of a pornographic nature to mean material that has either been issued an R18 classification certificate from the BBFC or material whose nature is such that it is reasonable to expect that, if it was submitted to the BBFC for a classification certificate, it would be issued an R18 classification certificate.[60]

4.96 Other material that has either been issued, or would be likely to be issued, an 18 classification certificate as a "sex work" by the BBFC will also be regarded by Ofcom as

[58] ICO guide to the UK GDPR
[59] ICO Age Appropriate Design Code (The Children's Code)
[60] See BBFC's definition of Sex works at 18.

restricted material of a pornographic nature, similar to the approach to R18 or R18-like material.[61]

4.97 When a VSP is considering whether its service has restricted material of a pornographic nature or not, it should have regard to the BBFC's definition of such material as *works whose primary purpose is sexual arousal or stimulation.*

4.98 We do not currently recommend or endorse specific technological tools or methods that a VSP provider should use to restrict access to pornographic material, though the chosen access control measure(s) should be effective in preventing access to that material for under-18s. We expect providers to stay informed of emerging technological developments and solutions for online safety and consider these as part of their ongoing assessment of the measures that are appropriate for their service.[62]

4.99 VSPs should seek to provide users with a clear understanding of the age verification method(s) that they are being asked to use on the service and, if more than one method is available, accurate information on the choice of those methods.

4.100 Ofcom would not consider the following forms of age verification to be appropriate protection measures for material of a pornographic nature:

- Self-declaration of date of birth or a 'tick box' system to confirm that the user is over the age of 18;
- General disclaimers asserting that all users should be deemed to be over the age of 18;
- Relying on age verification through online payment methods which may not require a person to be over 18, e.g. Debit, Solo or Electron cards or any other card where the card holder is not required to be over 18;
- Relying on publicly available sources or otherwise easily known information such as name, address and date of birth to verify the age of a user.

Protecting under-18s from material unsuitable for classification

4.101 Material which has either been determined not suitable for a classification certificate by the BBFC or material whose nature is such that it is reasonable to expect that it would not be suitable for a classification certificate (see Section 3) should be considered by VSPs as restricted material that has the most potential to harm the physical, mental or moral development of under-18s.[63]

4.102 We expect VSPs that feature this type of material to have in place a robust access control system that verifies age and prevents under-18s from accessing such material, in line with the expectations set out above for restricted material of a pornographic nature.

[61] See Section 368E (5) of the Act

[62] OSTIA – Online Safety Tech Industry Association is the industry body for UK organisations operating in the field of online safety. Its members are actively developing a range of solutions that a might VSP might consider implementing. https://ostia.org.uk

[63] There is no requirement for material being provided on a VSP to be classified by the BBFC

Protecting under-18s from other material that might impair their physical, mental or moral development

4.103 We recognise that there is a very broad range of restricted material which might impair the physical, mental or moral development of under-18s. In Section 3 of this guidance we have set out some examples of material that might impair under-18s, based on a literature review of existing evidence. VSPs may also wish to refer to the BBFC's age-based classification guidelines, which provide helpful information about the types of material that might be unsuitable for under-18s.[64]

4.104 VSP providers have to consider how proportionate their age assurance measures are in preventing access to under-18s, based on the harm that material might cause. We note again that the most harmful restricted material should be behind the strictest access controls. It may be appropriate for platforms to employ other protection measures in tandem with age assurance e.g. content ratings, parental control systems and other restricted mode settings to protect under-18s.

Age assurance and age-appropriateness for under-18s

4.105 In order for VSPs to manage material that may be age-inappropriate for specific age groups under the age of 18, they should seek to understand which age groups are using the service so they can ensure that material is age-appropriate, taking into account the different developmental needs and interests of under-18s. We use age-inappropriate here to refer to material which might harm the physical, mental or moral development of children based on their age group.

4.106 Whilst all platforms should be aiming for an age-appropriate experience for their users, we recognise that there are limits to how far VSPs are able to prevent age-inappropriate material from appearing on a service. We also acknowledge there are difficulties associated with verifying the actual age of under-18s using a service. Therefore, we share below a non-exhaustive list of considerations for effective age assurance that may guide VSPs in being able to protect the youngest and most vulnerable under-18s.

4.107 If a measure is aimed at estimating the age of an under-18 user this should be done in a way that appropriately safeguards children's personal data in line with the standards of the ICO's Children's Code, as well as the ICO's more general data protection requirements.[65]

Considerations for effective age assurance

4.108 VSPs may consider the following factors when establishing and operating age assurance systems:

[64] BBFC Classification Guidelines
[65] The ICO's Children's Code Hub provides a data protection code of practice for online services likely to be accessed by under-18s

a) It is important to assess, in a privacy preserving way, who is using the service. Higher risk services should make greater efforts to understand the age of their users.

b) VSPs should consider how reliable and accurate any age assurance method is and what level of confidence it provides, in relation to the risk, e.g. if a solution is not able to accurately distinguish between an adult and a child on a service that has the most harmful material, it is very unlikely to provide appropriate protection.

c) Age assurance measures that are easily integrated into existing platforms and avoid disrupting the user experience are likely to be more widely adopted and sustainable in the long term.

d) Some under-18s can provide false information to easily bypass age assurance measures, e.g. self-declaring to be 18 or over. VSPs should aim to have a robust and effective age assurance approach to account for and disincentivise this behaviour. Examples of this can range from neutral design of the date of birth request upon sign-up with no further chance to sign in if an underage declaration is made, to introducing hard identifiers or account verification for users who claim to be over 18.

e) VSPs should consider how different tools such as ratings and parental controls might interact with age assurance to provide greater confidence about the age of under-18 users. Trust-based measures such as parental controls may provide alternative and lower-risk forms of authentication and verification for under-18 users (see Parental Control systems below).

f) In certain circumstances, a provider may consider that using third-party age verification services, or compliance with a third-party age certification scheme may be a practical way to achieve a greater confidence level in the age of its users, especially if it is not feasible to develop in-house solutions. This may also be a way of complying with any technical standards in this area as they develop, such as PAS 1296, the Government's Digital Identity and Attributes Trust Framework or any other relevant standards.[66]

g) When considering age assurance solutions, VSPs should also seek to understand any potential exclusionary risks to children that might result from a particular type of measure, e.g. children in care may not be in an environment that is conducive for a parental consent solution, or the age of children with special educational needs may not be easily verifiable through behavioural or profiling methods.

[66] The UK Digital Identity and Attributes Trust Framework – which sets rules governing the future use of digital identities - was published in draft in February 2021.

Consultation questions

Question 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?

Question 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.

Parental control systems

*Provide for parental control systems in relation to **restricted material.***

4.109 Parental control systems allow an adult responsible for a person under the age of 18 a degree of control over what content the child can see or hear. Providers who offer services to under-18s should strongly consider having some form of parental control feature to support their overall protection measures for under-18s.

4.110 There are a range of parental control features that VSPs are able to design and implement. Some VSPs have systems which link accounts between child and parent or carer, thus giving the parent control over the type of content that their child can see (for example, through a 'restricted mode' setting), as well being able to restrict who can view the child's uploaded content.

4.111 Other features include restrictions on screen time, direct messaging and privacy. While VSP regulation focuses on video content, sophisticated parental control solutions that include features that apply to text and communication can enhance user protections by reflecting how under-18s use all the functionalities of a service.

4.112 Although parental control systems can also be applied at the network or device level, it is important that VSP providers consider the role they can play in promoting trust and safety through parental controls on their platforms. [67]

4.113 While VSPs can have flexibility in their approach to parental control tools, providers should consider the following factors:

a) Parental control systems can work most effectively where there is the establishment of a trust-based dynamic between the parent or carer, the under-18 user and the VSP.

[67] Parental control tools across networks and devices will often use some form of user-authentication and involve the parents' account, or the device linking to a child's account through use of the child's login credentials.

b) VSPs may consider the benefits of paired accounts between a parent or carer and an under-18 user, as well as the potential enhancement of safety through password-protection features on the service.

c) Parental control functions should not be easily circumventable by under-18s. Passwords or PINs required to unlock parental controls may be fixed, or might be in the form of one-time verification codes. We consider that one-time verification codes offer greater protection from under-18s potentially bypassing these systems.

d) Adults responsible for under-18s should be given only as much control as is necessary to protect them from harmful material, i.e. under-18s should be able to enjoy age-appropriate content and activity on a service without undue parental interference. This is especially important for older teenagers who should have a higher level of autonomy than younger users.

e) A parental control feature that can also verify that the user is a child is likely to increase a VSP's overall age assurance level. Providers should consider advancements taking place in parental consent technologies and whether a service might benefit from the early adoption of parental controls to increase age assurance and overall safety.

f) The use of parental controls should be informed by a VSP provider's understanding of the type of age-inappropriate videos that may be on their service, rather than just relying on limiting communication or interactivity features. This may be achieved by linking to any existing age rating system that the provider has in place. In the absence of an age-based rating system, it may be useful to consider existing UK classification systems, such as the BBFC age ratings or PEGI age ratings for interactive content, as a basis for restricting content under a parental control system.

g) The reliability and success of parental control systems will depend on the level of trust in the parent/carer-child relationship, as well as the availability of tools that may circumvent parental checks (e.g. VPNs, changing of login details).

Consultation question

Question 9: Do you have any comments on the draft guidance provided about parental control systems?

Complaints process

In relation to the implementation of the [flagging and reporting mechanisms and the explanations associated with them; systems for users to rate harmful material; age assurance systems; and parental control systems], 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.

4.114 VSP providers should have a complaints process that allows users to raise issues with the platform. Where providers choose to have such a process in place, the VSP Framework

suggests it should relate to the following measures: flagging and reporting mechanisms and the explanations associated with them; systems for users to rate harmful material; age assurance systems; and parental control systems.

4.115 Although the complaints process in the VSP Framework is limited to these measures, 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.

4.116 This complaints process should be available to both users and non-users of a VSP. For example, a parent or guardian may complain about parental controls failing to protect a child from restricted material, or to report an under-age user on a service they do not use.

Ensuring complaints processes are transparent, easy to use, and effective

Transparent

4.117 Providers are required to ensure their complaints process is transparent. Accordingly, it should be clear to anyone wishing to make a complaint how the process works from the outset, before they make a complaint. This might include providing: the information required from a complainant to make a complaint; the likely timeframe of the complaint process; potential outcomes of a complaint; and information on what communication the complainant can expect from the VSP provider throughout the process.

4.118 It is best practice for complainants to receive clear information about what will happen to their complaint once submitted, likely timeframes for resolution and communication about the ultimate outcome.

4.119 We expect responses to complaints to be appropriately timely as well as proportionate to the size and nature of the platform and the issue being complained about. We expect a platform to set its own internal timeframes for responses to complaints and to regularly review and record performance against these.

4.120 We recommend that data regarding the number of complaints and their outcomes be collected by the VSP provider to improve its complaints processes and its understanding of users' exposure to harmful content on the platform and how they are engaging with the protection measures. Such information might support Ofcom's monitoring of the VSP Regime (see Section 7).

4.121 The impartial dispute resolution procedure (as set out from 4.128) and how to access it should be communicated to the complainant at the conclusion of the complaints process.

Easy to use

4.122 The complaints process and information about it should be available to anyone accessing a VSP and be located in an easy-to-find area of the platform. The way information is presented may cause users to be disengaged from complaining. Therefore, we suggest that platforms should have regard to the same considerations regarding length and readability, as set out under the terms and conditions protection measures above, when providing information about the complaints process (as well as the outcome of any complaint).

4.123 VSP providers should consider the platform's user profile when designing a complaints procedure (see Section 5). We expect VSP providers to take an inclusive approach that takes account of the needs of vulnerable users. As explained above in paragraph 4.12, VSP providers might, for example, consider: using simple language, not overcrowding information, highlighting or emboldening key pieces of information, making sure the complaints process is easy to find and clearly highlighted to complainants, and making sure information is accessible, for example that it is readable by screen reader software.

4.124 The procedure itself should be in line with the way users typically engage with the VSP. For instance, it would not be appropriate if the only way someone could submit a complaint was via telephone.

Effective

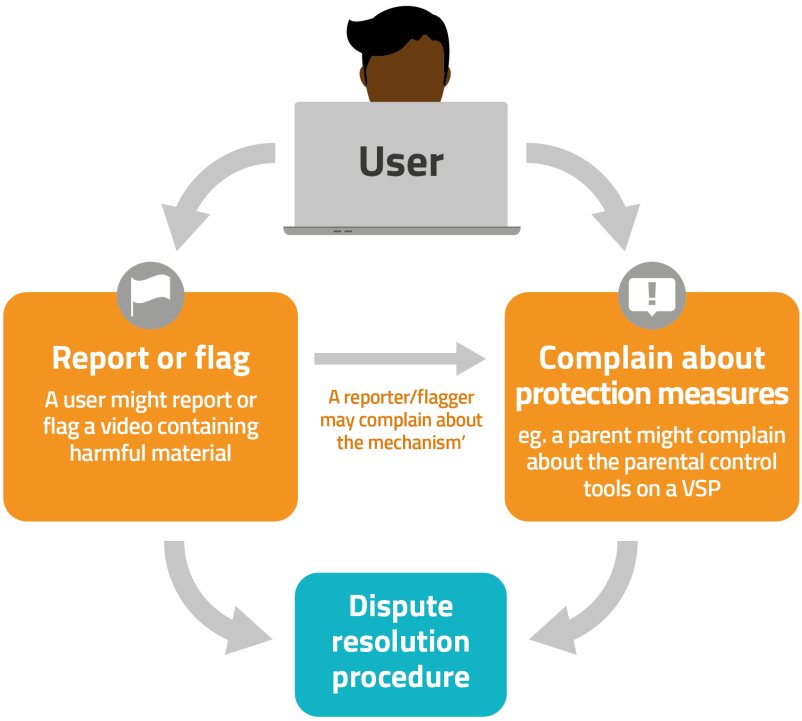
4.125 Measuring the effectiveness of a complaints process is challenging. Quantitative analysis of complaints figures may be useful to draw some conclusions but is likely to be impacted by myriad other factors. Indeed, increased complaints can sometimes reflect an improved process. One other quantitative indicator might be the proportion of complaints which result in escalation to a separate dispute resolution procedure (see below).

4.126 Potential methods of measuring effectiveness might include VSP providers: conducting user satisfaction surveys or seeking feedback from complainants; checking drop-off of complainants during the process and understanding reasons behind this; looking at data from social media insights; conducting reviews of the process with staff who deal with complaints; tracking complaint levels to inform knowledge of trends over time; and carrying out root cause analysis in instances where issues are identified. These potential methods are not an exhaustive list and it is for the VSP provider to ensure that the complaints process it has in place is effective.

Relationship with dispute resolution

4.127 It is important to note that the complaints process explained above is distinct from the dispute resolution procedure referred to below. Dispute resolution will typically, although not always, be relevant after an individual has exhausted the complaints process. We outline how a user might typically raise concerns and interact with these processes and procedures in Figure 4.1 below.

Figure 4.1: How a user might raise concerns with a VSP



Dispute resolution procedure

> *A person who provides a video-sharing platform service must provide for an impartial out-of-court procedure for the resolution of any dispute between a person using the service and the provider relating to the implementation of any measure set out in Schedule 15A, or a decision to take, or not to take, any such measure, but the provision of or use of this procedure must not affect the ability of a person using the service to bring a claim in civil proceedings.*

4.128 Under the VSP Framework providers are required to have a dispute resolution procedure in place that allows users to challenge decisions taken by the VSP provider and seek redress.[68]

4.129 This requirement covers all measures listed in the VSP Framework, including disputes about measures which are more directly related to the advertising-specific requirements (see 2.31).

4.130 The approach VSP providers take to implementing and operating a dispute resolution procedure may vary according to the nature of the service and its size (more information on these criteria are set out in Section 5). However, in order to fulfil the regulatory requirement, all VSP providers must have a dispute resolution procedure that is impartial.

[68] Section 368Z1(7) of the Act. The 2018 Directive refers to this procedure as an 'out-of-court redress mechanism'.

We also suggest that an effective dispute resolution procedure should be transparent, easy to use and fair, by which we mean that it should not unduly discriminate between users, introduce bias or result in inconsistent application.

4.131 The dispute resolution procedure must be available to users who are dissatisfied with the outcome of a decision made by the VSP provider about the application of protection measures. This could be, for instance, in relation to: a piece of content a user had reported using the provider's reporting or flagging mechanisms; or a decision by a VSP provider to remove a piece of content, suspend a user's account, or place sanctions on a user.

4.132 Outcomes of the procedure might result in, for example: the removal of content; sanctions against offending users; the reversal of wrongful content removal or sanctions; issuing an apology or correction; changes to processes or policies; or some other form of user redress, appropriate to correct the action in question.

4.133 If a dispute continues after a user has exhausted the dispute resolution procedure, VSP users have the ability to bring a claim in civil proceedings and the existence of the above processes and procedures must not affect that ability.

Ensuring dispute resolution procedures are impartial

4.134 In order to be considered impartial, a platform must be able to demonstrate procedural separation between its complaints or reporting processes and its dispute resolution procedure.

4.135 Ofcom considers that the most effective means of achieving impartiality that a VSP provider might take is to have in place an external, fully independent decision-making body or person, to meet this requirement.

4.136 Platforms may also choose to employ the services of an appropriately qualified third party or industry body for the mediation of disputes between a user and a provider. Ofcom is aware that this is an immature but developing sector in the UK and is supportive of innovation in this area to expand the range of online dispute resolution options available to platforms. However, VSP providers remain responsible for ensuring they provide a mechanism for the resolution of user disputes and ensuring that these are fit for purpose.

4.137 Ofcom acknowledges that designating an external, fully independent decision-making body or employing the services of a third party may not be a feasible model for all platforms and that other options might include a designated person or team internally, with responsibility for carrying out the dispute resolution procedure and reaching final decisions, that is procedurally separate from the complaints or reporting and flagging process.

4.138 In order to be impartial, the procedure should involve, at the very least, separate individuals dealing with the original complaint and the related dispute. The individuals should also be in separate teams, where the size of the platform allows for this.

4.139 This list is by no means exhaustive and it is for the VSP provider to ensure that the dispute resolution procedure that is in place meets the requirements of the regime.

Ensuring dispute resolution procedures are transparent, easy to use and fair

4.140 In addition to the statutory requirement that dispute resolution procedures are impartial, Ofcom considers it important that such mechanisms are also transparent, easy to use and fair. Transparency here covers both the dispute resolution procedure and the decisions and actions taken. This should involve:

a) clear signposting to users about how to appeal a decision;

b) information about what will happen once an appeal has been submitted, including how it will be considered and the anticipated timeframes for resolution;

c) clear communication with users during and after the dispute procedure, including an explanation of the outcome, the criteria against which the appeal was considered, and any actions or redress to be taken; and

d) retaining information about the number of appeals submitted and their outcomes, which may be requested by Ofcom. We would also encourage the publication of such information.

4.141 Dispute resolution procedures should be easy for all users to find and engage with, including provisions for vulnerable users (see 4.12). VSP providers should consider the platform's user profile when designing a dispute resolution procedure (see Section 5).

4.142 The procedure itself should be in line with the typical way users engage with the VSP. For instance, it would not be appropriate for the only way for a user to submit a dispute to be via telephone.

4.143 Enabling users to challenge decisions around the application of measures is important to ensure actions taken are fair. An effective dispute resolution procedure will help platforms to quality assure that they are enforcing their terms and conditions in a way that does not unduly discriminate between users, introduce bias or result in inconsistent application. Effective application should thereby mitigate the risk of over-takedown of content.

4.144 In instances where a VSP provider may have removed a piece of content or blocked a user from the service, it will be key that the dispute resolution procedure in place has clearly taken into account the importance of protecting users from serious harm and balanced it against the user's rights and legitimate interests.

4.145 Careful consideration should be given to disputes about videos containing news content. There may be instances where relevant harmful material or restricted material features as part of a news report, for example, and the inclusion of the harmful material is necessary for the purposes of informing or educating the audience. In these instances, the VSP provider should consider the context in which the potentially harmful material is presented, such as intention of the inclusion of the material as well as any information it has about the user uploading the material.

4.146 Similar consideration should be given to disputes concerning material on a VSP uploaded by a regulated broadcaster. Broadcast content is subject to stricter rules under the Broadcasting Code and therefore we would generally not expect this content (where it has

not been edited or presented in a materially different way) to raise an issue under the specified areas of harm in the VSP regime if it has already been complied for broadcast. Edits or clips of broadcast content however, may not retain the same contextual considerations as the original material. For some platforms it may be appropriate to consider an expedited process for the handling of disputes from broadcasters and other media outlets.

Consultation question

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

Media literacy tools and information

Provide tools and information for individuals using the service with the aim of improving their media literacy and raise awareness of the availability of such tools and information

4.147 Ofcom defines media literacy as "the ability to use, understand and create media and communications in a variety of contexts".[69] In the context of the VSP Framework the focus is on how media literacy has the potential to help ensure users of VSPs are protected from harmful material, both in the sense of creating or accessing such material, and its impact on them.

4.148 Media literacy can help protect users from harmful material broadly by encouraging safe use, critical understanding and responsible creation, so that users are better equipped to curate, understand and interpret their experiences when using a VSP. This can include building users' understanding of the safety features available on VSPs.

4.149 When designing and implementing this protection measure, VSP providers should:

a) Consider the specific tools and information needed to **improve users' media literacy** based on the nature of their service and the types of users on it;

b) Consider how they can **raise awareness** of tools and information to improve media literacy; and

c) Take steps to **understand the effectiveness** of the tools and information to improve media literacy on an ongoing basis.

4.150 VSP providers should also consider the five principles set out at paragraphs 4.8 to 4.14 when designing and implementing this protection measure:

- **Effective:** tools and information to improve media literacy should practically help users to protect themselves or others. VSP providers should consider the content contained in the tools and information and how they are designed and delivered, alongside the

[69] Information about Ofcom's media literacy activities

impact they are seeking to have on user skills or behaviours (see understanding effectiveness below at 4.154).

- **Easy to use:** tools and information to improve media literacy should be easy to locate, understand and engage with. This should take into account who they are aimed at, how and when they are communicated in the user journey and, if relevant, the device(s) that are likely to be used to access them. VSPs should have regard to the same considerations as set out under the terms and conditions protection measures in paragraphs 4.27 to 4.36 regarding length, readability, format, location, timing and promotion.
- **Fair and transparent:** tools and information to improve media literacy should place only a reasonable expectation on users to be responsible for taking steps to protect themselves or others from harmful material and should articulate the role the VSP plays in protecting its community of users from harmful material.
- **Evolving:** tools and information should be regularly reviewed and amended to reflect changing attitudes and behaviours, and new information sources as they become available.

4.151 Media literacy is also an important consideration in the design and implementation of other protection measures, so platforms should take a holistic approach to media literacy on their service beyond the provision of specific tools and information.

Improving users' media literacy

4.152 VSPs should consider the specific information and tools needed to improve users' media literacy based on the nature of their service and types of users on it. This could include, but is not limited to, exploring ways to minimise the impact of harmful exposure on users and providing tools and information that help users develop and exercise skills in:

a) Using the VSP in a safe and secure manner to maximise opportunities and minimize risks. This might include:

   i) Understanding the safety features available on the VSP; the benefits of using them; and how to use them.

   ii) Understanding the mechanisms available on the VSP to support users if they are exposed to, or have created, harmful material; the benefits of using them; and how to use them.

   iii) Awareness of the potential benefits and possible risks of using the VSP.

b) Critically understanding the VSP and the content available on it in order to make informed choices and best manage use. This might include:

   i) Awareness of how content is delivered to users and, if relevant, how content recommendations are informed and made.

   ii) Recognising different types of content available on the VSP, such as advertising and its different forms.

iii) Understanding how the VSP is funded and how the user contributes to it (e.g. if fees are paid or if user data is used to inform advertising).

c) Creating content responsibly to support positive experiences for themselves and others. This might include:

i) Understanding what is and isn't acceptable content on the VSP.

ii) Understanding the potential consequences of creating unacceptable content.

Raising awareness

4.153 VSPs will need to consider how to raise awareness of tools and information to improve media literacy as part of the overall design of their user experience and journey. We recognise that the points at which it might be most effective to raise awareness may need to be tested and iterated, but we expect VSPs to consider raising awareness in the following broad settings:

a) **During the registration process or relatively soon afterwards.**

VSPs should consider how to best make users aware of the availability of tools and information to improve media literacy at an early point in the user journey. VSPs could consider whether default settings that opt for the safest setting are appropriate, particularly for under-18s.

b) **At regular intervals as users participate on the platform.**

VSPs should consider how awareness can be raised throughout the user journey. Providers could consider whether regular check-ins or proactive reminders are appropriate, particularly if default settings have been changed, or whether ongoing visibility of some tools and information on a landing page is appropriate.

c) **Before, during and after viewing of harmful material.**

VSPs should consider how they can raise awareness of resources that help users respond appropriately to harmful material close to the point of engagement with it. VSPs could consider signposting users to these resources if they search for terms likely to generate harmful material; placing warning labels on harmful content that advise users of the nature of it and directing them to helpful resources; and signposting users who flag or report harmful material to helpful resources.

Understanding effectiveness

4.154 We encourage VSPs to take steps to evaluate the effectiveness of the tools and information they provide to improve users' media literacy on an ongoing basis. This could include understanding users' awareness of and engagement with the tools and information, and whether engagement resulted in the intended outcome. These findings can then be used to improve the tools and information provided if appropriate.

4.155 We recommend that VSPs also consider best practice from, and partnerships with, other organisations to ensure they are creating and delivering tools and information to improve

media literacy in the most effective way. One way to learn from others is through media literacy networks that aim to increase collaboration, information-sharing and debate to improve media literacy.[70]

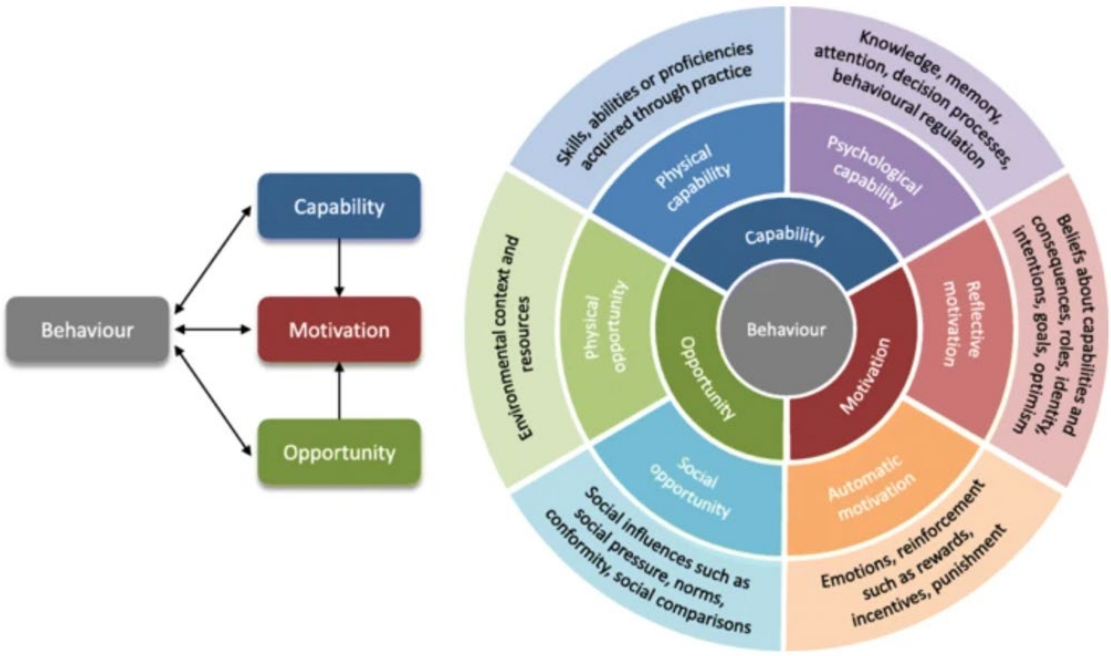
Designing effective protection measures - understanding user behaviour

4.156 Media literacy is one of a number of key considerations in the design and implementation of all protection measures, alongside other factors which influence user behaviour and engagement. For example, the physical location and appearance of tools and information, and where they are located on a screen. Therefore, we encourage VSPs to assess whether their users can engage with all measures in a way that protects them from harm.

4.157 One way in which VSP providers can achieve this is by systematically analysing whether there may be behavioural factors that could limit the effectiveness of their measures. For example, as mentioned above, users may lack awareness of the measures, or the measures themselves may be designed in such a way that makes engagement difficult.

4.158 The COM-B model is one framework that can help identify how the three elements that influence human behaviour (capability, opportunity, and motivation) can act as a barrier to user engagement.[71] This is illustrated in Figure 4.2 below.

Figure 4.2: The COM-B model



Source: Michie, S., Van Stralen, M.M. and West, R., 2011.

[70] Ofcom have a media literacy network called the Making Sense Of Media Network.
[71] Michie, S., Van Stralen, M.M. and West, R., 2011. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implementation science, 6(1), pp.1-12.

4.159 For example, the COM-B model can help with understanding how the physical or psychological capabilities of a user might impact their propensity to engage with reporting and flagging mechanisms, whether the physical environment (opportunity) encourages or discourages users to engage with terms and conditions, or how a user's motivation may deter them from seeking media literacy guidance or information.

4.160 An analysis of these barriers combined with other evidence about user behaviour (e.g. consumer research) can then be used to develop and test interventions that could be effective at reducing these barriers.

Consultation question

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

5. Determining which measures are appropriate

5.1 Under the VSP Framework, VSP providers are required to determine whether it is appropriate to take a particular measure to protect users from harmful material according to whether it is practicable and proportionate to do so, taking into account:

a) the size and nature of the video-sharing platform service;

b) the nature of the material in question;

c) the harm the material in question may cause;

d) the characteristics of the category of persons to be protected (for example, under-18s);

e) the rights and legitimate interests at stake, including those of the person providing the video-sharing platform service and the persons having created or uploaded the material, as well as the general public interest;

f) any other measures which have been taken or are to be taken.[72]

5.2 In addition, when determining whether a measure is appropriate, VSP providers must apply the principle that restricted material that has the most potential to harm the physical, mental or moral development of under-18s must be subject to the strictest access control measures.[73]

5.3 Ofcom is required to provide guidance on the measures listed in the VSP Framework, as well as their implementation. In our view, the practicable and proportionate criteria are useful not just for providers to determine *which* measures to take but also, in some circumstances, *how* to take those measures to achieve the required protections. Guidance in this section can be used alongside the general guidance on the implementation of measures set out in Section 4 above.

5.4 Decisions about which measures to take and how to implement them are related to the risk of harm to users on a platform. A low risk of harm is, generally, likely to require fewer or less sophisticated protection measures compared to a VSP with a greater risk of harm.

5.5 In Section 6 we encourage all VSPs to conduct some form of risk management process and suggest a framework for this. When providing information about some of the practicable and proportionate criteria below, we do so with these risk assessments in mind.

5.6 It is important to note that the criteria are covered in the order in which they are listed in the VSP Framework. This does not represent an order of importance as all the criteria should be considered in the round.

---

[72] Section 368Z1 (4) of the Act. The Act contains an additional criterion relevant to advertising: in relation to adverts that are not marketed, sold or arranged by a person providing a video-sharing platform service, the fact that the provider exercises limited control over such communications.

[73] Section 368Z1 (5) of the Act

Size of the platform

5.7 A VSP's size may be determined through a range of metrics, including reach of the platform and volume of content. Resources are also an important consideration.

5.8 However, there are no particular thresholds above or below which particular measures should be taken. Instead, providers should consider these metrics in the round alongside the other practicable and proportionate criteria.

Reach and volume of content

5.9 A higher volume of content (i.e. videos available on a service) is likely to increase the risk of harmful material being available; and the more users a service has (i.e. the number of people who might see or engage with that content), the more people are at risk of encountering harmful material. Additionally, high numbers of users or a large volume of content can present moderation challenges which can increase the risk to users.

5.10 Relevant metrics when assessing the reach of a platform might include: the number of users; the level of engagement with the service (e.g. the number of unique page visits to a platform's website, or the time spent on the website); and the number of accounts (active or otherwise).

5.11 VSP providers should also be aware of the reach of their service beyond the boundaries of their own platform. A platform's popularity might mean that content is shared between platforms or re-broadcast across other mediums.

5.12 Relevant metrics when considering the volume of content could include the number, combined length, or combined size of videos uploaded to a service, either within a set timeframe or in total. The number of channels and the average length or average number of videos uploaded by users at any given time might also be relevant.

Resources

5.13 Understanding the resources of a VSP provider is relevant. Financial and other considerations, such as staffing and the size of the entity providing the VSP service, may affect the viability of taking particular measures and the way in which they are implemented. We understand that some of the most sophisticated measures set out in Section 4 may only be practicable and proportionate for the largest platforms.

5.14 However, cost and resources cannot be considered in isolation when determining whether a measure is practicable and proportionate. What is practicable and proportionate must be considered in the round and weighed against the risk of harm to users on a platform. Lack of resources or profits alone may not be sufficient justification for not taking a protection measure. This is particularly so for material that has the most potential to harm and is required to be subject to the strictest access control measures.

Nature of the service

5.15 Providers should consider the type of service they provide. The functionality that determines how users engage with the service (i.e. how they watch, upload and share videos) and how it is operated (i.e. funding and business model) are important considerations in looking at the nature of a service.

Functionality of the service

5.16 VSPs enable users to upload and share videos with members of the public in a wide variety of ways. For example, videos may be accessible to every user of a service, or the service may provide the functionality for users to determine who can see their content. Videos may be permanently available or they may be ephemeral (viewable only for a time-limited period). Videos may also be streamed in real-time or pre-recorded and uploaded. These different functionalities present different moderation challenges, and platforms may need to strengthen other measures as a result.

5.17 Services also present content to users in a number of ways. For example, on some services, the primary route of discovering content is through an active search. On others, content is continually fed, video after video, to a user whose engagement and selection of content is much more passive.

5.18 Continual feeds may present a higher risk of users encountering content they do not want to see (including harmful material) and so platforms should consider whether stricter measures, such as a pre-warning message before the content plays, need to be applied. VSP providers must understand how their algorithms feed content to users in order to understand the risk profile of the platform and should consider user safety when designing and reviewing such processes (see Section 6 – Embedding a safety-first approach).

Business and operating models

5.19 Business models might impact the type of users or content on a platform (e.g. deliberately targeting business-to business customers) which could affect the risk profile of a service (see 'The characteristics of the category of persons to be protected' below). Business models might also impact the functionality of the service (e.g. offering a white-listed version of the service which might have a different level of oversight).[74]

5.20 Some VSPs are funded through subscriptions, as opposed to purely advertising-based models. Subscription services may have a different risk profile compared to an open, free platform because of the different type of engagement with their users.

5.21 For some VSPs, the type of content they host is central to their business model. For example, services which deliberately position themselves as anti-censorship. The nature of

[74] "White-listing" refers to providing a product or service to a customer but removing all of the provider's branding. It can often include a greater level of autonomy in the way the product or service is used.

these services may require stricter access control measures (i.e. age assurance or parental controls) than services which are not focused on such material.

The nature of the content in question and the harm it may cause

5.22 Different types of content carry different levels of risk of harm. Whether a service specialises in particular genres of videos may be relevant, for example. Further, providers should seek to understand the prevalence of potentially harmful content on their platform. Potential applicable metrics to consider here could be the volume and types of content reported by users.

5.23 VSP providers should have regard to Section 3 where we set out the types of content likely to be considered as relevant harmful material and restricted material, as well as considering the harm this material may cause to users.

5.24 We also encourage providers to conduct their own research into the harm particular content might cause to users, particularly if the content in question is distinct from that set out in Section 3.

5.25 It is important to stress again that the VSP Framework includes the principle that restricted material that has the most potential to harm the physical, mental or moral development of under-18s (such as pornography or gratuitous depictions of violence) must be subject to the strictest access control measures.[75]

The characteristics of the category of persons to be protected

5.26 It is important for a VSP provider to have a good understanding of who the users on their platform are, particularly whether there is a high proportion of children in the user base. Platforms will be better able to tailor their protection measures to the relevant audience using this information.

5.27 Understanding VSP users' age is likely to be one of the most relevant ways in understanding the category of persons to be protected and therefore which measures are practicable and proportionate. Knowing the age of users will also aid the implementation of these measures (e.g. understanding whether a simplified version of terms and conditions aimed at under-18s would be appropriate).

5.28 In determining whether a measure is practicable and proportionate to take, VSP providers should also consider users of their platform who may be vulnerable (see 4.12).

5.29 When attempting to understand the characteristics of a typical user on a VSP platform, providers should have regard to data protection law requirements, including the ICO's Age Appropriate Design Code and its principle of data minimisation.[76]

[75] See section 368Z1 (5) of the Act
[76] ICO Children's Code Hub

The rights and legitimate interests at stake, including those of the service provider and the users having created or uploaded the content and the general public interest

Interests of the users

5.30 When considering the proportionality of taking and implementing any particular measure providers must take into account its potential impact on the rights and legitimate interests of users, particularly those who engage with the service as uploaders and sharers of content. For example, where legitimate interests may be impacted, protection measures and their implementation need to be proportionate to the harm the provider is seeking to address. This is likely to be especially relevant in the context of any measures that could apply to block or restrict a user from uploading content or which require a user's content to be removed.

5.31 Providers should also take into account the rights and legitimate interests of users when designing and operating their systems and procedures to: ensure they do not unduly discriminate between users or introduce bias; that their actions are transparent and consistent; and that they provide opportunities for users to challenge content-related decisions (see Complaints Process and Dispute Resolution Procedure in Section 4).

5.32 Other rights and legitimate interests of users may be covered by other regulatory regimes. Users' rights to privacy and data protection are covered by GDPR. These interests might be relevant when taking, implementing or assessing the effectiveness of measures. For example, where personal data is used to assess or improve the effectiveness of measures, this should be separated from any processing for commercial purposes.

Interests of the service provider

5.33 Where a VSP provider is taking appropriate measures to protect its users from harmful content, we do not expect that the choice of protection measures should disproportionately limit or impede a VSP provider from conducting its business as it chooses. However, in assessing a provider's compliance with the VSP Framework, including which measures have or have not been taken, Ofcom must balance the rights of the service provider against the objectives of the VSP Framework to protect users from harmful material.

5.34 Careful consideration will be given to situations where it appears that a provider has not taken a particular protection measure which, in Ofcom's view, would otherwise have been appropriate for it to have taken, because it prioritised its commercial interests over the need to protect users from harmful material. This will be especially relevant where harm has occurred and can be directly linked to a failure to take the measure or it appears to Ofcom that there is a high risk of such harm occurring as a result of the measure not being taken.

General public interest

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.

Consultation question

Question 12: Do you have any comments on the draft guidance 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?

6. Additional steps to protect users

6.1 In this document we have provided guidance on the measures in the VSP Framework and how providers can implement those measures effectively on their platforms to secure the required protections for users.

6.2 We consider that there are additional steps platforms could take to strengthen the protection for users. These are related to the protection measures but are not necessarily a requirement of the VSP Framework. We provide information about some of these here, noting that many of the examples given are elements of existing good industry practice. We consider that platforms taking these steps are more likely to be in a position to secure appropriate protections for users from harmful material.

6.3 We also strongly encourage providers to conduct risk assessments when determining the measures required to protect users on their platforms. We provide guidance on risk management processes and how they might be applied to VSPs in this section.

6.4 In order to be effective, we anticipate an element of information and data collection regarding the protection measures, to be used by the platforms themselves and requested by Ofcom from time to time. We suggest the types of information and data which might be considered useful in this section.

Embedding a safety-first approach

6.5 VSP providers are encouraged to take a safety-first approach and to design their service with this in mind. This means considering the needs of users in all decisions about the service to develop a culture of safety.

6.6 Some VSP providers have told us their trust and safety teams work closely with other teams throughout the development of a new product or feature, allowing them to influence safety at an early stage and evaluate whether new products, functionalities and services put users at risk of harm. We consider collaboration between safety teams and other teams across a VSP provider's organisation to be a positive step, including engagement with senior executives. We also recognise that clear accountability for safety at senior levels can promote greater consideration of user safety across an organisation.

6.7 While we encourage VSP providers to ensure all teams within their organisation consider user safety, providers may also choose to have an individual or team in place who are specifically responsible for user safety.

6.8 This role or team's responsibilities could include; overseeing internal governance processes; driving internal debate and dialogue on safety; supporting enforcement of safety policies and practices; embedding internal expertise on significant harms; obtaining external specialist advice on harms; facilitating partnerships with organisations tackling known illegal harms, such as CSEA and terrorism; and ensuring a rapid response team is in place to deal with most egregious or illegal harms that require immediate action.

6.9 We understand that some online content providers operate "intelligence desks" which aim to anticipate content trends and investigate harmful behaviours that are currently undetected by a platform's autodetection mechanisms. This can involve platforms investigating methods being used by bad actors to circumvent protection measures, such as the use of codewords for promoting racial hatred.

6.10 VSP providers may also choose to establish boards or groups to help inform user safety. These can include external experts and/or members of the community. They can be used to inform decisions or improve features on the platform; promote or protect the interests of marginalized groups; or to strengthen relationships with regional regulators or law enforcement agencies. We welcome providers gaining insights from users and experts in this way.

External engagement

6.11 External engagement plays an important role for many providers in supporting the protection of users from harmful material. Here we set out some examples that have been provided through our engagement with industry. This is not an exhaustive list and where we refer to specific organisations this should not be viewed as an endorsement by Ofcom.

Third party content moderation

6.12 VSP providers may choose to use third party content moderation to assist with the identification, removal and reporting of harmful content.

6.13 **ActiveFence** identifies and tracks harmful content online, including disinformation, child sexual abuse, hate speech and terror content. **Thorn** builds technology to protect children from sexual abuse. Thorn's 'Safer' tool helps small- and medium-sized services find, remove and report child sexual abuse material (CSAM).

6.14 While we welcome and encourage the use of external experts in this way, ultimately the responsibility for ensuring that protection measures are effective at protecting users from harm lies with the VSP provider, even where such third parties are involved.

6.15 Third party content moderation can be used by VSPs to help assess the prevalence of certain types of content on their platform, as well as evaluating the effectiveness of their moderation systems. They may also be used to report illegal content to relevant law enforcement, such as reporting CSAM to the **National Center for Missing & Exploited Children (NCMEC), the Internet Watch Foundation (IWF)** or **NCA-CEOP** (National Crime Agency).

Charities, NGOs and harms experts

6.16 We encourage VSPs providers to work with charities, NGOs and academics to bring specialist insight and knowledge into the development and implementation of policies and procedures. Drawing on such expertise will help providers develop their understanding of

particular topics or issues, including what help, support and protections users (and staff) may need.

6.17 For example, **the Samaritans** have worked with providers to create guidelines on how sites and platforms hosting user generated content can manage self-harm and suicide content and keep vulnerable users safe online.

6.18 Providers may consider seeking the views of specialist charities and NGOs when developing its policies and practices, including the **NSPCC**,[77] **Holocaust Educational Trust,**[78] **Tell MAMA**[79] and **Beat.**[80]

6.19 For criminal content such as terrorism and CSAM, we suggest all VSPs seek specialist advice from organisations such as the **IWF, NCMEC** or **Tech Against Terrorism** (see Section 3) and encourage those with a higher risk profile for this type of content to explore more formal partnerships or memberships.

6.20 Further information on the UK safety tech sector that may be helpful in supporting providers in developing their approach to protecting under-18s, can be found in the report on the protection of children online published alongside this guidance.

6.21 Collaborations may also be useful when considering media literacy. For example, a number of providers partner with **Internet Matters**, which offers advice and resources to parents and families to help keep children safe online. Other services include **Childnet, ThinkUKnow** and the **NSPCC's Net Aware.** Providers can also learn from others through membership of media literacy networks that aim to increase collaboration, information-sharing and debate to improve media literacy. We encourage VSPs to join Ofcom's Making Sense Of Media Network.

6.22 Ofcom will also continue to engage with the broad range of experts in the online space to gain insights and inform our approach to VSP regulation.

Assessing and managing risk

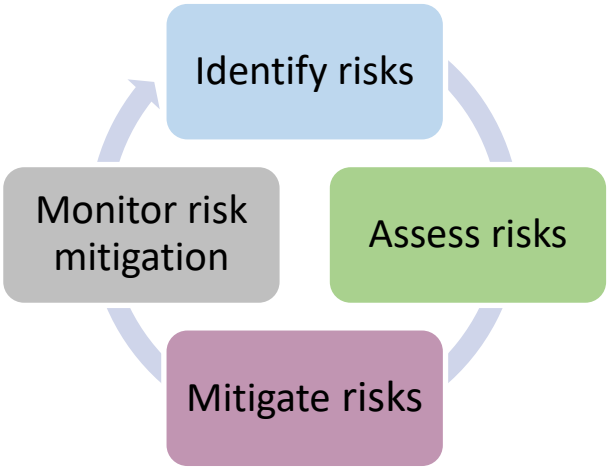
6.23 In order to determine which measures are appropriate for protecting users on their platforms, we strongly encourage VSP providers to put a process in place to assess and manage risk. Although this is not a requirement under the VSP Framework we think it is a clear way for providers to document the decisions they have taken when determining which measures are appropriate to protect users from harmful material. We note that the assessment and management of risk will likely form part of the future online harms regime and is also part of the European Commission's Proposal for a Regulation on a Single Market For Digital Services (Digital Services Act).[81]

[77] The National Society for the Prevention of Cruelty to Children
[78] The Holocaust Educational Trust
[79] Tell MAMA supports victims of anti-Muslim hate and measures and monitors anti-Muslim incidents.
[80] Beat is a UK eating disorder charity.
[81] The Digital Services Act: ensuring a safe and accountable online environment

6.24 Below we provide guidance on how VSPs might conduct this process. This is presented as a step-by-step guide to support providers, but we recognise that individual platforms might approach risk in different ways.

Risk management process

6.25 The basic steps of a risk management process are **identifying**, **assessing**, planning for and **mitigating** any risks, and **monitoring** to ensure that risks are appropriately mitigated.



6.26 Risk management is a dynamic process and so risks should be reviewed regularly, and providers should consider whether the risks have been appropriately mitigated. It is good practice to document this process, clearly illustrating the steps taken by the VSP provider in considering the risks to users on the platform and how this has influenced the decisions about which measures to take.

6.27 Some VSP providers will already have mature risk management frameworks and processes, while the assessment and management of risk may be new to other VSP providers. Risk management can be undertaken by platforms of any size. Risk management is a framework to consider the actions that need to be taken in assessing and managing risk. It is important so that VSP providers can satisfy themselves that they have adequately considered the risks of harmful material on their platform and whether the measures they have taken are appropriate to protect users.

6.28 We expect that discussions with providers about how they assess and manage the risk of harmful material on their platforms will form part of our supervisory engagement and will promote greater use of risk management frameworks as the VSP regime develops. Ofcom may also ask about risk management as part of any enforcement activity, for example when assessing whether a provider has failed to take and implement a measure which we consider to be appropriate.

Identify the risk of harmful material on the platform

6.29 We encourage providers to carefully consider the risks to users that could arise from the characteristics of their platform. Some of the practicable and proportionate criteria covered in Section 5 may be useful for identifying these risks.

6.30 The nature of the service and the nature of the content carried on the platform may make particular risks more likely to arise. For example, the following may carry different levels of risks to users:

- The different ways videos are shared (e.g. users being able to determine who can see their content; videos being permanently available or time-limited; and videos which are pre-recorded and uploaded or streamed in real-time).[82]
- The way content is found and consumed (e.g. actively searched for or presented as continual feeds, where the selection of content is more passive)
- The type of content a particular service might focus on (e.g. extremist views or pornography)

6.31 In addition, evidence of risk could come from historical experience of harmful material on the platform, external research, consulting with third party safety organisations and engagement with their users.

6.32 Understanding how a platform has identified the risks of harmful material on their service is likely to be an area to be explored in supervisory discussions. We encourage VSP providers to collect relevant data, for example via the reporting and flagging mechanisms they may have in place.

6.33 As part of this, it could be useful for platforms to have the ability to categorise different types of harmful material, with reference to the material set out in Section 3. Such categorisation will aid the understanding of a platform's own risk profile and so help determine whether additional measures are practicable and proportionate. Such information may be useful as part of compliance monitoring.

Assess the risks of harmful material on the platform

6.34 Having identified the relevant risks to users encountering harmful material on the platform, it is important that VSPs consider the impact of that harmful material.

6.35 Again, some of the practicable and proportionate criteria set out in Section 5 may provide a useful framework for considering the severity and seriousness of the risk. For example, the size and the user base of the platform could play an important role in considering the likelihood of harm occurring and the level of the impact. In particular, where under-18s form a significant proportion of the user base, they would be a key factor for consideration

[82] The Full Government Response to the Online Harms White Paper noted that "a service is likely to be higher risk if it has features such as: allowing all users - including children - to live-stream themselves".

in the risk profile because they are a clear category of persons to be protected under the VSP Framework (see Section 3).

6.36 Understanding the risk profile of its own service will help a VSP when considering the appropriateness of protection measures. For example, a VSP with a low risk profile is, all other things being equal, likely to require fewer or less sophisticated protection measures compared to a VSP with a high risk profile. As noted above, we would encourage VSPs to document their risk profile.

Consider whether existing protection measures adequately mitigate the risk of harmful material

6.37 Once a VSP provider understands their service's unique risk profile, we would expect them to consider whether existing protection measures on the platform adequately protect users from harmful content. VSP providers should consider a user's journey through their service and how they interact with different measures at different points. In particular, some questions that VSP providers may want to ask themselves are the following:

a) What measures are currently in place to protect users from harmful material?

b) How effective are they to protect users and how are we monitoring the effectiveness of those measures over time, with what indicators and data, to understand when measures need to be improved?

c) What is the internal governance process surrounding the taking of measures and monitoring their effectiveness?

6.38 We would expect a VSP provider to specifically consider all measures listed in Section 4 as well as any other (non-specified) measures that the VSP has implemented.[83]

6.39 After considering the impact of existing protection measures individually, a provider might then make an overall assessment using all available information of how the measures taken in the round are protecting its users against harmful material. If a provider identifies a need to add additional measures or make improvements to existing measures, they could:

a) assess what additional mitigation options are available;

b) assess the likely effectiveness, practicability and proportionality of the additional mitigation options; and

c) then select what additional mitigations to add on the basis of this assessment.

6.40 An important part of the exercise under this step is understanding the effectiveness of existing measures. VSP providers should collect information about the usage and impact of their protection measures and review that against their own risk profile. Further detail on information collection can be found below.

[83] For example, modifications to an algorithm that minimises the visibility of harmful material.

Continue to monitor effectiveness of protection measures to manage risks and protect users

6.41 Good risk management is an ongoing process. The risks of harmful material on a VSP are likely to be constantly changing, so VSP providers should regularly monitor the effectiveness of their protection measures to protect their users. We expect VSPs may need to regularly track data that indicates the level of harmful material, risk of harm and the impact of individual measures on these outcomes. If existing measures are not effective at providing mitigation against risks of harm, providers should consider whether further measures, or an extension of existing measures, are needed.

Consultation question

Question 13: Do you have any comments on the draft guidance around assessing and managing risk?

Measuring effectiveness

6.42 Whether or not a platform puts in place a risk management process as encouraged, or follows the framework above, Ofcom considers that measuring effectiveness of protection measures is vital for platforms to understand how well they are working to protect users from harmful content. The collection of data and information is an important aspect of this. Such information may be helpful to Ofcom in assessing compliance.

6.43 We strongly encourage VSP providers to collect proportionate information about the measures which have been taken and implemented on their platforms and how effective these measures are at protecting users from harmful material. Such information is likely to support a VSP provider's risk assessment and related decisions about taking further measures or strengthening existing ones.

6.44 This information might include quantitative metrics on user interaction with protection measures, as well as data indicating the prevalence of harmful material, where feasible. Examples of quantitative metrics collected by platforms to test the effectiveness of their protection measures could include:

a) Volume of reported harmful material (from users, trusted flaggers, and automated systems)

b) Accuracy of systems which identify and remove harmful material

c) The number of violations of terms and conditions relating to harmful material

d) The number of content-related decisions appealed, as well as the proportion of successful appeals (as an indicator of wrongful removals)

e) User awareness of and engagement with protection measures

6.45 In addition, some platforms also collect information on the views that the reported or removed content received. The number of views received by the violating content is an

important factor in assessing the effectiveness of measures, as it gives an indication of the prevalence of harm and how many people have seen the content prior to it being actioned.

6.46 We recognise that the relevant context must be considered along with numerical indicators. For example, an increase in the volume of content removed may be due to an increase in violating content being uploaded by users, an increase in monitoring by the VSP or by a change in policy by the VSP provider. Given this, qualitative indicators should also be considered when evaluating the effectiveness of protection measures.

7. Ofcom's approach to monitoring and enforcement

7.1 Ofcom has a duty to take steps to ensure that VSP providers comply with their requirements under the VSP Framework, which include taking appropriate measures to protect users from harmful material.[84] One of the ways we will achieve this is through monitoring and enforcement of the VSP Regime.

7.2 In this section we set out some of the ways we currently expect to monitor compliance. We also provide some information on how Ofcom will approach enforcement.

Monitoring

7.3 Monitoring VSP providers' compliance with obligations is an important aspect of ensuring an effective regulatory regime - it helps us understand what industry is doing and ensures we take effective action where we may have concerns either about the whole of the industry, or about individual platforms.

7.4 Through our monitoring we will work collaboratively with industry to help establish common understanding about how protection measures and their implementation can appropriately protect users.

7.5 To help us identify how providers are implementing the VSP Regime, we are likely to use a combination of the below tools. Our approach to monitoring will likely evolve as we move through the different stages of the VSP regime, and in response to intelligence gathered.

7.6 The response we take to any compliance concerns raised through our monitoring will vary depending on the nature of the concern raised and we set out more detail on this in the enforcement section below.

Informal engagement including supervision

7.7 We will engage with VSP providers as appropriate to ensure they understand their new obligations and how they can comply with them. Through our informal engagement we will seek to ensure that providers have implemented appropriate measures to reduce and mitigate the risk of harmful material. We expect to continue this engagement throughout the regime but anticipate having more focused supervisory activity in the early stages of the regime to ensure providers are taking steps to meet requirements.

7.8 The purpose of this informal engagement or supervisory activity with platforms is not to identify compliance concerns so that we can take enforcement action, but to ensure that providers understand the obligations and have the tools they need to comply with them.

[84] Sections 368X (1); 368Y (1); and 368Z1 (1) (a) and (b) of the Act.

We will be clear with platforms as to the nature of our engagement and if we identify concerns that may lead to more formal enforcement action.

7.9 We will also continue to engage with other stakeholders including charities, NGOs and harms experts to gather wider expertise and insights to inform our regulatory activities.

Information gathering

7.10 Ofcom has broad statutory powers to request information from VSP providers for the purposes of fulfilling its functions as the regulator.[85] This includes information needed for monitoring providers' compliance with the VSP Framework.

7.11 For the purposes of monitoring, such information requests are likely to focus on gathering information to help us to understand how platforms are ensuring compliance with the VSP Framework, including consideration of:

a) Which measures a platform has in place and, where relevant, which measures the provider has decided would not be appropriate to put in place.

b) How those measures are implemented (including their effectiveness at protecting users from harmful material).

c) Any risk management (or similar) processes, which inform a provider's decisions about the measures in place on the platform.

7.12 For (b) we have suggested ways a platform might measure effectiveness above at paragraph 6.42. Guidance on risk management processes is set out at 6.23

7.13 In addition to information collection for the purposes of monitoring compliance, Ofcom has information gathering powers for the purpose of producing and publishing reports[86] about:

- Steps taken by providers of video-sharing platform services to comply with their duties, including appropriate measures to inform viewers about advertising and impartial dispute procedures.
- The measures taken by providers to protect users and the ways in which such measures are implemented.
- The systems adopted by providers for the reporting, flagging or rating of material and the handling of complaints or the resolution of disputes relating to the service.

7.14 When issuing information requests Ofcom will clearly set out the purpose for which it has been sent and the reasons we consider we need the information to fulfil our duties. We will also ensure these are justifiable, proportionate and fair, and will have regard to the need to exclude from publication, so far as practicable, matters which are confidential.[87]

[85] Section 368Z10(1) of the Act. Section 368Z10(2) also provides Ofcom the power to request information from a person who is not necessarily a VSP provider.
[86] Section 368Z11 of the Act.
[87] In accordance with sections 36810(5) and 368Z11(2) of the Act

Accessing VSPs

7.15 We may access VSP platforms in scope of the regulation from time to time. This may be as part of a proactive assessment of a VSP provider's compliance with the requirement to take measures to protect users from harmful material, or in response to specific concerns about particular harms on a platform.[88]

Complaints

7.16 We will accept complaints from users of VSPs via an online webform from summer 2021. However, we will encourage users to complain first to the platform in question about harmful material so that these videos can be removed or restricted as appropriate. Users can complain about these processes.

7.17 Ofcom will not respond to or adjudicate on individual complaints. Instead, Ofcom will review the information holistically, for example by analysing trends in the volume of complaints across platforms or by reference to particular harms. This will help Ofcom to identify cross-industry concerns and potential issues with compliance and determine the appropriate regulatory response.

7.18 Ofcom will also welcome complaints from other interested stakeholders, including charities and tech safety groups. However, we reiterate that concerns should be flagged to the platform in the first instance so that providers have the opportunity to take appropriate action.

7.19 Ofcom will consider concerns raised by other EEA regulators where they have identified issues relating to appropriate measures taken by UK-established VSPs, as well as complaints relating to cases where there is risk of serious harm to users.

Enforcement

7.20 Where we identify compliance concerns through our monitoring, we will consider whether it is appropriate and proportionate to take enforcement action, or some other action, to help protect users from harm.

7.21 Ofcom has the power to take enforcement action where a VSP provider has failed to:

a) Notify Ofcom that it is in scope;

b) Pay any fee that is required by Ofcom;

c) Cooperate and/or comply with a statutory information request;

d) Provide for an impartial out of court procedure for the resolution of disputes;

e) Inform viewers about advertising that the provider does not control but is aware of; and/or

[88] Ofcom will always act in accordance with the law, including avoiding unlawful interception and unlawful surveillance.

f) Take or implement appropriate measures to protect users against harm;

7.22 Where (f) above is concerned, there are likely to be relevant indicators which lead us to consider whether there are compliance concerns with a platform. These could include:

- A high prevalence of harmful material and/or that material being easily accessible on the platform;
- Delays in responding to reports, flags or complaints about harmful material;
- A high volume of complaints, to the provider or to Ofcom, about harmful material on the platform or about the measures taken by a provider; or
- A lack of engagement with Ofcom.

7.23 Although Ofcom has the power to take formal enforcement action if a VSP provider breaches its obligations, where appropriate we will attempt to work with providers informally to try to resolve compliance concerns before using our formal enforcement powers. For example, we may informally request information from platforms to help us to understand the issue in question and the steps the VSP provider has taken to address the issue and comply with its obligations.

7.24 We would be much more likely to take immediate formal enforcement action where we suspect a potential breach if it appears that:

- The type of harmful material appearing on a platform is that which has the potential to cause serious and significant harm to users and/or that harmful material is easily accessible or appearing for a prolonged period;
- The provider appears to be taking no action in response to being alerted to the presence of harmful material appearing on its platform;
- The platform has demonstrated a poor record of harmful material appearing on the platform, particularly if Ofcom has previously engaged with the provider about the same or similar issues; or
- The provider is not engaging with Ofcom about the steps it is taking to protect users.

7.25 In the event of a compliance concern regarding taking or implementing appropriate measures to protect users against harm, the questions Ofcom will likely want to consider include:

a) Which measures the platform has taken to protect users from harmful material.

b) Whether those measures have been implemented in such a way as to effectively protect users from harmful material.

c) Whether it would have been practicable and proportionate for the VSP to have taken any of the other measures set out under the VSP Framework or for them to be implemented in another way.

7.26 If Ofcom decides that formal enforcement action is necessary, we will investigate the issue to determine: if there has been a breach; if a sanction should be imposed; and if so, what sanction is appropriate. The subject of the investigation will be given the opportunity to make representations before a final decision is made. If Ofcom does find that a breach has occurred, we have the power to impose various sanctions. For example we could impose a

financial penalty, require the VSP provider to take specified actions, and/or suspend or restrict the service.

7.27 Any enforcement action will be taken in line with our Enforcement Guidelines.[89] Ofcom also has published information on the setting of financial penalties in our Penalty Guidelines. The considerations set out in that guidance will be applied to any VSP enforcement notice where relevant.

[89] These Guidelines are currently being updated to reflect new powers Ofcom has been given. Ofcom will update this Guidance with the new Enforcement Guidelines when they are published.