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: 6 October 2021

Contents

Section

1. Overview

What this guidance covers

This guidance is intended to support UK-established video-sharing platform (VSP) providers in understanding their regulatory obligations. The statutory framework applying to these services came into force on 1 November 2020 (referred to as "the VSP Framework" or "the VSP Regime"). 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;
- explains the measures platforms can take to protect users from harmful material, and provides guidance on how to implement those measures effectively;
- explains the practicable and proportionate criteria providers must consider when determining which measures to take. These criteria include the individual characteristics of VSPs and the rights and legitimate interests of users;
- encourages providers to establish risk management processes to 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". Guidance on specific requirements relating to advertising standards will be published separately.

The VSP Framework sets out measures for providers to take, as appropriate, to protect users from harmful material. These include:

- 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

VSP providers are not required to take all these measures, but should determine whether it is appropriate to take a particular measure, according to whether it is practicable and proportionate to do so, considering factors 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. Measures must be implemented in a way that protects users from harmful material.

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 allowing companies to innovate in the systems and processes they use to keep users safe. 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. While the legislation sets out measures that providers are required, as appropriate, to take and implement, providers may also choose to take other actions to protect users.

Below we have listed the main protection-related requirements on providers and our expectations regarding compliance.

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’s terms and conditions require 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 pornography or material unsuitable for classification, or where there is a significant risk of under-18s encountering such material.
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 ensure that VSP providers comply with these requirements, which we will carry out by working with industry and other stakeholders, and through our monitoring and enforcement activities.

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. The VSP Regime is about platform’s safety systems and processes, not about regulating individual videos, however evidence of a prevalence of harmful material 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 about which protection measures to use.

Along with engagement with providers themselves, we expect to inform our understanding of whether users are being effectively protected, for example 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.

Ofcom’s understanding of the measures companies have taken, and what constitutes effective protection for users, will develop over time. Both platforms themselves, and the risk of harm on those platforms, will evolve. Providers may also choose to take other actions to protect users that are not set out in the legislation and we will want to understand these too. We will work with providers to ensure services are clear about what is expected of them. To support this, we will publish annually our priorities for UK-established VSPs, and when necessary, we will update this guidance.

We will also publish an Annual VSP Report to provide transparency on the measures providers are taking to protect users and to report on progress made against the priorities we set out.

Where Ofcom has concerns, we will act in accordance with our Enforcement Guidelines. Where appropriate, we will generally seek 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. Sanctions could include an enforcement notification requiring the VSP provider to take specified actions, and/or impose a financial penalty. Ultimately, we also have the power to suspend or restrict a service in cases involving the most serious non-compliance. We will use our enforcement tools proportionately and where we consider the evidence shows they are justified, having regard to the right to freedom of expression while increasing user safety.

More information about Ofcom's plan and approach to VSP regulation, as well as our current annual priorities can be found on the Ofcom website.

2. Introduction

2.1 Video-sharing platforms, or "VSPs", are subject to regulations in the UK. In this section we provide an overview of the 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 we expect it to be expanded further in future following Government's announcement that Ofcom will be appointed as the online safety 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, but we are independent from Government.

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, and 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 may be appropriate for providers to take 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, came 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. 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 5% 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 is not empowered to resolve individual complaints about items of content. Instead, the Framework focuses 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 have regard to the right to freedom of expression.

[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 Guidelines will be updated to reflect new powers Ofcom has been given. Ofcom will update this guidance with the new Enforcement Guidelines when they are published, following consultation.

Regulatory context

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 a 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 May 2021, the Government published its draft Online Safety Bill, confirming Ofcom as the regulator of the future online safety regime. At the same time the Government confirmed its intention to eventually repeal Part 4B of the Act, meaning the regulation of UK-established video-sharing platforms will be superseded by new legislation following the commencement of the online safety 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 safety regime as described by Government in the draft Online Safety Bill.

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), and the Financial Conduct Authority (FCA) through the Digital Regulators Cooperation Forum (DRCF) to support regulatory coordination in online services and cooperate on areas of mutual importance.[12]

[11] Draft Online Safety Bill

[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 came into force on 6 April 2021. Existing UK-established VSP providers were required to notify their services by 6 May 2021. Ofcom has published a list of notified services. Services commencing after 6 April are required to make an advance notification to Ofcom of their intention to provide a service. See our guidance 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 has the power to set and charge VSP providers an annual fee from April 2022. We will consult providers in good time before setting up any fees regime.[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 the 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 consulted 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. The VSP advertising guidance, once finalised, will be available on the Ofcom website.

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 may be appropriate for platform providers to take in order to protect users from the harmful material defined

[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.

[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".

above. It is important to note 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 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”. Whether a measure is appropriate for the purposes of protecting users from harmful material, must be determined by whether it is practicable and proportionate. 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 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.

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

How to use this guidance

2.36 This guidance is designed to help VSP providers understand what is expected of them under the VSP Framework, as well as other steps Ofcom encourages providers to consider. 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 must take such of the measures set out in the VSP Framework, as appropriate to secure the required protections. **Section 4 of this guidance sets out these measures.**

b) Where measures are taken, they must be implemented effectively so 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) **Section 5** sets out guidance around the statutory requirement for VSP providers to have an **impartial dispute resolution procedure.**

d) Providers must determine which of the measures to take according to whether it is practicable and proportionate for a measure to be taken.

   i) **Section 6 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 7 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.

2.37 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. This document provides guidance on the framework; it should not be viewed as a set of compulsory steps. Providers should use the guidance to support their considerations on which measures are appropriate for their service to secure the required protections.

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

2.39 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 8 has further details about Ofcom's approach to monitoring compliance and enforcement.**

3. Harmful material

3.1 The VSP Framework requires VSP providers to take the measures (set out at 2.33 above) appropriate 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 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.

3.4 We expect providers to be aware of these 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 7).

3.5 Ofcom intends to work with VSP providers to understand how they have taken account of the risk of different types of Harmful Material on their platform and how they have

[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.

incorporated the relevant legislative definitions in the Act into terms and conditions. Some providers may also choose to adopt acceptable use policies which go beyond the definitions of the legislation; this remains at the discretion of providers. More information about our supervisory approach to engagement can be found in Section 8.

3.6 We will also keep our guidance under review and consider what further information or practical guidance might be useful to support the development of policies on restricted material and relevant harmful material as the regime matures.

Restricted material

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

3.7 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.

3.8 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 an R18 classification certificate.[25]

Material unsuitable for classification

3.9 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]

3.10 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]

3.11 While the **VSP Framework does not require videos on VSPs to obtain a BBFC classification**, providers will need to regularly consult BBFC guidelines to understand the type of material that is likely to be unsuitable for classification. BBFC guidelines can help services stay informed of UK consumer expectations about the type of content standards relevant to them and draw from specialist expertise to inform their own community guidelines.

[25] See BBFC Classification Guidelines on R18 material

[26] Section 368E (3) (a) and (b) of the Act

[27] See BBFC Classification Guidelines.

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

3.12 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.

3.13 The legislation does not specify 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 when drafting terms and conditions or acceptable use policies and determining which other measures it may be appropriate to take, to protect under-18s from material that might impair the physical, mental or moral development:

a) **Pornography;**[29]

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.

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

3.15 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.

3.16 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]

[28] See UEL Report for more examples and information on the academic evidence to support these categories of harm.

[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

3.17 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.

3.18 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 emerging new harms and changing attitudes.

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

3.20 Overall, while VSPs must ensure the measures they take are effective in protecting under-18s, they should also have regard to the benefits children and young people derive from using the service e.g. to acquire knowledge, connect with others, and seek enjoyment and self-expression.[33]

Relevant harmful material

3.21 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

3.22 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 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]

3.23 VSP providers will need to ensure their policies take into account both 'incitement to violence' and 'incitement to hatred'. With respect to the latter, 'hatred' should be understood as referring to a feeling of animosity or rejection with regard to a person or a group of persons targeting one or more of the protected characteristics listed above. VSP providers should ensure that any 'incitement to hatred' policy is not limited to only incitements to violence.

3.24 When determining the appropriateness of the measures VSPs take, Ofcom will have regard to relevant case law on freedom of expression, which includes the case law of the

[31] The Cap Code: The UK Code of Non-Broadcast Advertising and Direct & Promotional Marketing

[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.

European Court of Human Rights (ECHR).[35] In September 2020, the ECHR published a factsheet summarising some of its cases on incitement to hatred, which may be helpful to providers.

3.25 VSP providers 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 negative stereotypes that exist and the social context. In assessing whether content is 'likely' to incite violence or hatred amongst general users and against the targeted group or person with particular protected characteristics, providers should pay attention to the potential effect of such content.

3.26 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 online hate. The examples in the report of industry approaches may be relevant for VSP providers considering their own approach to hateful content.[36]

3.27 In Section 4, on page 26, we have provided some suggestions about how providers might consider drawing up and effectively implementing terms and conditions which prohibit the uploading of material likely to incite violence or hatred.

Material the inclusion of which would be a criminal offence

3.28 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.

3.29 VSP providers should establish internal protocols to escalate and action this type of material as soon as it is identified. 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).[37]

3.30 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.

3.31 It is ultimately up to VSP providers themselves to determine what types of material might constitute a criminal offence if included in the service, seeking their own legal advice as

[35] ECHR cases can be found here.

[36] Ofcom, A Report by The Alan Turing Institute on Understanding online hate, 2021.

[37] For the avoidance of doubt, platforms should not send potentially illegal content to Ofcom.

appropriate.[38] Providers may also wish to consider working in partnership with others with relevant expertise. (See External Engagement in Section 6 for organisations which may be able to support providers on these matters). Providers should also ensure that relevant terms and conditions are not defined any more narrowly than the scope of the criminal offences elaborated below.

Terrorism

3.32 The VSP Framework refers to Article 5 of Directive (EU) 2017/541 of the European Parliament and of the Council of 15 March 2017 on combating terrorism, which requires prohibition of any statement published directly or indirectly encouraging terrorism or that is likely to be understood as such. It is irrelevant whether any person is in fact encouraged or induced by the statement to commit, prepare or instigate a terrorist act.

3.33 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

3.34 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").

3.35 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. 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. Throughout the rest of this document we refer to this as child sexual abuse material (CSAM).[39]

3.36 VSP providers should be aware that youth-produced sexual material (sometimes referred to as ‘sexting’, ‘nudes’, or ‘nude selfies’) also constitutes CSAM and should be dealt with under policies relating to combating this material.

3.37 Online sexual grooming, while not covered under the VSP Regime, is an illegal activity that can lead to the production of CSAM.[40] As a result, providers may wish to consider policies to address these wider illegal harms, as doing so may reduce the risk of CSAM being uploaded to their service.

3.38 In relation to CSAM, VSPs may wish to familiarise themselves with the Memorandum of Understanding agreed between the Crown Prosecution Service (CPS) and the Association

[38] Criminal offences, as defined in the legislation, referred to in the definition of relevant harmful material in s.368Z1

[39] ‘Child pornography’ reflects the language used within the legal framework for this regime, but this is no longer commonly used in the UK. Children's advocates tend to use the term child sexual abuse material.

[40] Note that online grooming is illegal under UK and EU law under the CSEA directive.

of Chief Police Officers (ACPO), which may be helpful when planning what to do if CSAM content is flagged on their platforms.

Racism and Xenophobia

3.39 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.

3.40 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.

3.41 Also included are offences related to publicly condoning, denying or grossly trivializing crimes of genocide, crimes against humanity, war crimes and other specified crimes,[41] 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.

[41] 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.[42] Guidance on measures that are most relevant to ensuring advertising standards are met will be published separately on the Ofcom website.[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 6 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
- Age assurance systems
- Parental controls systems
- Complaints process
- 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.[45] However, one of the aims of effective implementation should be to prevent

[42] Ofcom deems a user as anyone able to access, view or upload content on a platform, not just those who have an account. In the context of certain protection measures, a user could also be a parent or guardian.

[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. Draft guidance on the regulation of advertising on video-sharing platforms can be found here: https://www.ofcom.org.uk/consultations-and-statements/category-1/regulation-of-advertising-on-vsp. This guidance will be available in its final form later in 2021.

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

[45] For completeness, we note that some platforms have chosen to adopt proactive monitoring technologies, although the VSP framework does not impose an obligation on platforms to pro-actively monitor for harmful material.

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 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.[46] Where we think effective protection of users is unlikely to be achieved without a specific approach, we say so.

Five principles to support implementation

4.7 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**.

4.8 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:

[46] 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, including updates to the service itself, to ensure that they remain appropriate for their intended purpose.

Measuring effectiveness

4.9 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. It is not a requirement to gather information for this purpose but any such information that a provider does collect 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

4.10 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

4.11 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 6).

4.12 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 unwarranted 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 Complaints Processes at 4.143 and Dispute Resolution in Section 5).

Considering user behaviour in designing and implementing measures

4.13 As well as considering the five principles which support implementation, we encourage VSP providers to assess whether their users can engage with all measures in a way that protects them from harm.

4.14 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. The COM-B model is one framework that can help identify barriers to user engagement and we provide more information about this at 4.167.

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.

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.*

4.17 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.

4.18 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.

4.19 One way to implement this measure could be to introduce a step into the upload process to prompt users to consider whether they need to notify the platform of restricted material.[47] This prompt could include information to help users assess their content correctly, for example, an explanation of what constitutes restricted material or a reminder of the platform's terms and conditions, including potential sanctions.

4.20 Providers should also consider what action they take to protect under-18s in response to being made aware of videos containing restricted material. **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 rating systems (see 4.82 – 4.103) or employing access control measures, such as age assurance or parental control tools which are discussed below from 4.104.

4.22 Where there is material permitted on the platform that may not be age-appropriate for all users, the provider should consider how implementation of this measure can provide an age-appropriate experience for different age-groups of users to protect those who are most at risk (i.e. material that may impair the physical, mental or moral development of younger age groups may not carry the same risk for older teenagers). For example, providers may consider restricting access to content using age-ratings.

4.23 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 implement other measures to protect under-18s from videos containing restricted material, such as having appropriately robust age assurance systems in place (see 4.104 – 4.127).

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.***

4.24 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. It also refers to material the inclusion of which would be a criminal offence under laws relating to terrorism;[48] child sexual abuse material;[49] and racism and xenophobia.[50]

[47] Research by the Behavioural Insights Team (BIT), a global social purpose company that generates and applies behavioural insights to inform public policy, has found that presenting users during the sign up process with 'just in time' pop-up textbox explanations, which explain how the user's personal information may be used by the company, can lead to increased understanding of privacy policies. This type of approach could also be used to remind users of key terms and conditions when uploading content.

[48] The VSP Framework refers to Article 5 of Directive (EU) 2017/541 of the European Parliament and of the Council of 15 March 2017 on combating terrorism.

[49] The VSP Framework refers to the CSEA Directive.

[50] 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.

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

4.26 **Ofcom considers this measure to be 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.**

4.27 Providers are encouraged to ensure that their terms and conditions are clear about how they will deal with illegal content reported to the platform. For example, we know that many platforms report illegal content to law enforcement agencies.

Recommendations for good practice regarding terms and conditions prohibiting material likely incite to violence or hatred

Ofcom recognises that whether any particular type of material is likely to cause incitement to violence or hatred is difficult to judge. In Section 3, we provide guidance on the meaning of incitement to violence or hatred. Below we offer some suggestions on how platforms might draw up and effectively implement terms and conditions which prohibit the uploading of material likely to incite violence or hatred.

VSP providers who take this measure should be sufficiently clear that any material likely to incite violence or hatred is expressly prohibited on their platforms.

Providers should also be clear in their terms and conditions the groups against which incitement to violence or hatred will be prohibited on the platform. Providers should consider reflecting the protected characteristics listed in the definition of relevant harmful material under the VSP Framework. We note that platforms may choose to include a wider set of grounds where incitement to violence or hatred will be prohibited on the platform.[51]

The definition of relevant harmful material refers both to incitement to violence and incitement to hatred and providers will need to ensure they take both into account in setting any terms and conditions. Terms and conditions should be sufficiently clear to enable users to understand what content is prohibited.

In particular, terms and conditions should be clear that 'hatred' refers to a feeling of animosity or rejection with regard to a person or a group of persons, aimed at one or more protected characteristics. Providers could provide examples or share case studies of ‘incitement to violence or hatred’ to help users understand what content is likely to be prohibited, or set out examples of contextual factors that may be relevant to the platform’s assessment of whether or not the content overall is likely to incite hatred (e.g. whether challenge is included, the genre of the content, and any external circumstances affecting how it is likely to be understood).

VSP providers might also wish to clearly communicate how the severity of the material likely to incite to violence or hatred affects the prioritisation of content for moderation and corresponding proportionate sanctions.

Ensuring terms and conditions are easy to use

4.28 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

[51] As noted in Section 3, the definition of Relevant Harmful Material includes incitement to 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.

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

4.29 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.

4.30 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 and providing clear explanations of key concepts and requirements. This is particularly so given that around 15% of UK adults have poor literacy skills. [52]

4.31 Providers might want to consider ways in which the formatting of the service's terms and conditions can improve readability. For example, one study by the Behavioural Insights Team (BIT) found that using summary bullet-points with icons illustrating key terms or a Q&A summary, where terms are presented in a question and answer format, led to increased understanding of terms and conditions.[53]

4.32 Further, platforms that 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 understanding from younger users.[54] Platforms should also consider whether there are specific aspects of their terms and conditions that are particularly important for younger users to understand, for instance how to report or flag content.

4.33 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.[55]

4.34 Ofcom recognises that terms and conditions may need to cover a wide range of circumstances and behaviours. However, we consider that it is nonetheless possible to draft legally binding terms and conditions that may be easily understood by users or to provide an additional, more user-friendly version of terms and conditions, alongside the full legal document. Third-party organisations can provide resources and advice on how to draft user-friendly terms and conditions, for example, the Children's Commissioner's 'Growing up digital' report and BIT's 'Best practice guide'.

Location, format, timing and promotion

4.35 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

[52] See the National Literacy Trust's webpage on adult literacy levels.

[53] Improving consumer understanding of contractual terms and privacy policies: evidence-based actions for businesses, The Behavioural Insights Team.

[54] "Growing Up Digital", Children's Commissioner, January 2017.

[55] For example, CHANGE's 'How To Make Information Accessible' guide.

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.

4.36 Ofcom understands that some platforms use additional documents to explain their terms and conditions, for example explanatory blog posts or articles. We encourage platforms to ensure that terms and conditions are explained consistently in all communications with users, particularly when updating or reviewing terms and conditions and other explanatory documents. We suggest that providers making use of additional explanatory documents ensure that all rules and terms are also consolidated in one place on the site to make it easy for users to locate them.

4.37 We also consider that the method of accessing terms and conditions is important for 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.

4.38 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. 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. For more guidance on raising awareness of a platform's policies and safety resources, see paragraph 4.162.

4.39 It might be useful for providers to consider research into the presentation of terms and conditions when developing their own. One such example of this is BIT's best practice guide on presenting terms and conditions to consumers. [56] This is aimed at improving the number of people who open terms and conditions, read and understand them.

4.40 In BIT's guide, one technique that was found to lead to an increase in the number of people who open terms and conditions was telling users how long it would take to read the terms and conditions.[57] BIT also found that telling users when it was the last opportunity for them to review the privacy policy before signing up to a service led to an increase in the number of people who opened the policy.[58] They found that showing users the full text of terms and conditions within a scrollable text box (instead of requiring a click to view them) led to increased understanding of terms and conditions.

[56] Improving consumer understanding of contractual terms and privacy policies: evidence-based actions for businesses, The Behavioural Insights Team.

[57] We note that the BIT only tested this on terms and conditions with a reading time of less than 5 minutes. Further research would be required to determine whether the same approach continues to be effective for terms and conditions with longer with reading times.

[58] We note that BIT tested this technique – of informing users that it is their last chance to view the privacy policy – during the registration process rather than in the context of uploading content. VSP providers could consider extending this approach to inform users when it is their last chance to review the community guidelines before uploading a video, but this was not tested by BIT.

Ensuring terms and conditions are effective

4.41 Terms and conditions need to be implemented in a way that meets 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.

4.42 We encourage providers to assess how effectively they have implemented their terms and conditions. This might include, for example, evaluating users' engagement with terms and conditions by collecting data on the number of users clicking to open them or the amount of time users spend reading terms and conditions. We also encourage providers to assess adherence to and/or understanding of their terms and conditions.

Enforcement and sanctions

4.43 To work effectively, 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.

4.44 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.

4.45 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 VSP providers also block IP addresses.

Moderation

4.46 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).

4.47 Ofcom expects any moderation techniques used to be subject to regular quality assurance processes. This is so that they remain effective at enforcing the terms and conditions of a platform and, ultimately, provide effective protection to users from harmful material, while also seeking to avoid blocking or taking down legitimate content that is wrongly flagged as harmful. Ofcom would expect the accuracy of any machine-learning moderation techniques to be routinely checked by human moderators.

4.48 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 6). Some platforms 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.

4.49 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.

4.50 We encourage VSP providers to collect information to assess the effectiveness of any moderation techniques to support any risk management system they have in place. For more guidance on collecting information to measure effectiveness and on an approach for assessing and managing risk see Section 7. Ofcom may use this information to support its compliance monitoring (see Section 8).

Ensuring the implementation of terms and conditions is fair and transparent

4.51 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 need to be clear to all users. We encourage providers to consider setting out clearly in their terms and conditions all material prohibited on the platform and all potential sanctions.

4.52 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.

4.53 Where actions such as content removal, blocks or bans are taken, it is important users have the ability to challenge these decisions. This is vital to ensure that users (including users who are regulated broadcasters or other media outlets) are not unreasonably impacted by unwarranted takedown of content. Careful consideration should also be given to claims which involve videos containing news content and we explain this further in paragraph 5.21 below.

4.54 The dispute resolution procedure discussed in Section 5 is an important requirement under the VSP Framework and will help VSP providers ensure moderation and enforcement decisions are fair. Terms and conditions about harmful material should clearly signpost these functions and any other routes of appeal available.

Ensuring terms and conditions continually evolve

4.55 We expect VSP providers to keep terms and conditions under review and make changes where necessary. Examples of potential triggers for a further review of terms and conditions (in addition to regular, ongoing reviews) could include technological developments, new research, or changes to the online environment, to platform features or to relevant legislation.

4.56 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.

4.57 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.

4.58 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.

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;*
>
> *(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)*

4.59 **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.**

4.60 Reporting and flagging mechanisms enable viewers 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".[59]

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

4.62 Platforms could also consider using "trusted flaggers" to identify harmful content in a more robust way. For example, VSPs may want to prioritise reports from certain users (e.g. specialist charities) to highlight potentially harmful content faster. Should they do so, it is for the VSP to decide what the purpose of a trusted flagger is and who trusted flaggers are.

4.63 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

[59] "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).

does not follow that they must necessarily establish and operate a system to explain the effect of that report, though we note this second measure is likely to be good practice.

4.64 We recognise that the extent to which user reporting is effective can depend on the nature of the harm and the age of the user. For instance, research highlights that only a small proportion of children who are aware of reporting and flagging tools have ever used them to report content.[60]

4.65 Platforms should ensure that reporting mechanisms do not inadvertently disincentivise reporting of content. For instance, requiring an email address could discourage users from reporting illegal content.[61]

4.66 Ofcom considers that reporting and flagging mechanisms should be available to anyone able to view content on a platform, and not limited to those logged in via an account.

Reporting and flagging mechanisms should be easy to use

4.67 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.

4.68 VSP providers with a high number of under-18s using their service should ensure that their reporting and flagging mechanisms are designed so they can be easily used by children.

4.69 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.

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

4.70 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.

4.71 Platforms should explain clearly whether or not users can expect to be notified that action has been taken on their content as a result of a flag or report. In order not to prejudice any criminal investigation, it may be appropriate for platforms to provide that outcomes for illegal content it has reported to the police will not be proactively notified to the uploading user.

[60] Ofcom, Children and parents: media use and attitudes report 2020/21

[61] Reviewing Child Sexual Abuse Material Reporting Functions on Popular Functions, *Canadian Centre for Child Protection*, 2021.

4.72 Where a provider gives explanations about the effect of flagging and reporting, we encourage them 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.

4.73 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.

4.74 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.[62] 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 when they have viewed harmful material.

Responses to reports or flags should be appropriately timely

4.75 Responses to flags should be appropriately timely, proportionate to the size, nature and risk profile of the platform (see Section 6). We encourage platforms to set internal timeframes for responding to harmful material reported or flagged, and to review performance against these regularly. Such information might support Ofcom in assessing the implementation of this measure (see Section 8). Ofcom would expect such performance to improve over time.

4.76 Providers are also encouraged to prioritise responses to the most harmful material. Ofcom expects platforms to take the swiftest action in response to reports about the most harmful categories of content. For example, by having an expedited process for the handling of reports about terrorist material. This is particularly important in periods of heightened terror threats. In addition, providers should consider prioritising responses to viral content which has a significant number of views in a short space of time.

4.77 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.

4.78 Platforms should have the systems in place to inform users who wish to know how their report has been dealt with. 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.

4.79 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.

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

For example, by referring to the title of the video or a unique reference number for the report/flag.

Reports and flags should be effective

4.80 Reporting and flagging mechanisms can be effectively supported by internal escalation processes. 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.

4.81 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.

Systems for viewers to rate harmful material

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

4.82 This measure, as drafted in the legislation, 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.

4.83 VSPs may also operate ratings systems in which the ratings are generated by the platform or the uploader. We recognise that the measure in Schedule 15A only refers to systems for *viewers* to rate harmful material. We therefore discuss viewer ratings systems in this section and cover broader types of ratings systems below.

4.84 Platforms might allow viewers to apply ratings to content. If enough viewers believe that a rating should be applied this might be added without intervention from the platform being required. Alternatively, viewers may challenge or change ratings set by the platform or uploader where they think they are inappropriately applied. Platforms may also use these rating systems to test or improve the algorithms or other mechanisms which recommend content to users.

4.85 “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.[63]

4.86 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 from other types of rating system and we would also strongly

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

caution against its use to determine whether content reported or flagged as CSAM has been correctly identified.

4.87 For any system in which ratings could be changed by viewers without intervention from the platform, providers would need to be confident that their systems were not vulnerable to misuse and abuse.

4.88 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 as with other ratings systems we would not expect them to provide adequate protection when used in isolation from other measures.

Uploader and platform generated rating systems

4.89 Platforms may also operate systems where content ratings are applied by the platform or the uploader.[64]

4.90 At para 4.20, we suggest that, where a platform has terms and conditions requiring uploaders to notify them if a video contains restricted material, the provider should then take the additional step, as appropriate, of either informing viewers where a video contains restricted material or restricting access to it by under-18s.

4.91 Uploader rating systems could enable uploaders to notify the platform that a video contains restricted material. Uploaders could be asked to rate content being uploaded, according to a set of parameters set by the platform.

4.92 Where uploaders notify the platform that a piece of content contains restricted material, the platform could then use this information to apply a rating to the content. Ratings can take many forms - they may be age-based or may simply mean marking material as 'mature' (see paragraph 4.95 below).

4.93 Ratings, whether applied by the uploader or platform, or both, can be used to create an age-appropriate experience for users. For example, either by restricting access to material rated as unsuitable for certain ages, or by applying warnings in advance of the video being played.

4.94 This is a fast-evolving space and we expect that platforms will innovate to create new and effective systems in which the platform, viewer and uploader play different roles. Whichever is involved in applying ratings, we expect VSPs to implement appropriate ratings systems that are transparent and easy to use and are suited to working alongside other measures.

[64] These types of ratings system are not included in the appropriate measures listed in Schedule 15A, but platforms may wish to consider implementing these systems to support their other measures.

Different types of rating systems may be appropriate

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

4.96 Some platforms might consider having more sophisticated 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 artificial intelligence (AI) or machine-learning. VSPs with a large proportion of under-18 users may favour sophisticated systems in order to differentiate between the content that is appropriate at different stages of a child’s development.

4.97 Where VSPs are not implementing their own ratings systems, we suggest they may wish to incorporate established classification frameworks through partnerships with existing ratings bodies, such as BBFC[65] and VSC Rating Board (PEGI).[66]

4.98 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.

Ensuring rating systems are transparent and easy to use

4.99 Where systems to rate content are in place, explanations about which ratings apply to different types of content should be made explicit. Clear examples should be considered and providers should have regard to the suggestions about length, readability, location, format and timing at 4.29-4.40.

4.100 If platforms are using ratings systems as a way of ensuring they are notified that a video contains restricted material, they should make it clear to users what constitutes restricted material in accordance with the platform’s terms and conditions (with providers having regard to our guidance in Section 3).

Rating systems work well alongside other protection measures

4.101 Where platforms use access control measures such as age assurance or parental controls (see below), these can be tied to a rating system. Combining with parental controls may provide a mechanism for ensuring that under-18s cannot access material that a parent or caregiver determines to be age inappropriate, in line with the platform’s content rating boundaries.

4.102 Effective use of age assurance techniques can help platforms to understand the different age groups of under-18s using their service and therefore put them in a better position to

[65] BBFC Classification Guidelines
[66] VSC PEGI Ratings

deliver age-appropriate material. The use of age-based content ratings may also support this by signposting age-inappropriate content to viewers.

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

Age assurance systems

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

4.104 **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.[67] The term may also be used to refer to the level of confidence that a platform has in the age of its users.

4.105 Examples of age assurance cover a range of potential methods, from users self-declaring their date of birth to age estimation techniques such as 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.[68]

a) **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. Examples of age verification may include:

   i) Hard identifiers (passport scans, credit details, driving license, electoral roll information);

   ii) 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 that minimise the need for repeat authentication or verification.

b) **Age estimation** refers to methods that can estimate or infer a person's age, usually by algorithmic means. This may include, but is not limited to, the following techniques:

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

   ii) 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;

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

[67] 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)

[68] For more detail on how parental verification tools may support better age assurance, refer to Parental Control Systems below.

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

c) **Account confirmation** through the use of parental control software and mechanisms allows for existing account holders to confirm the age of a user. The benefits and limitations of account confirmation as part of a trusted parent/carer and child relationship is discussed in more detail in the Parental Control Systems section below.

d) **Self-declaration** is where a user states their age or date of birth but offers no further evidence to confirm the information. This is a measure that is easy to bypass by the user, who is able to enter the minimum age that allows access to a service that may carry age-inappropriate or harmful material for the actual age of the user.

4.106 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, and select an approach (or approaches) that is proportionate to risk, 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.

4.107 When considering 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[69] and The Age Appropriate Design Code.[70]

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

4.108 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.

4.109 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, VSP providers 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.

4.110 Ofcom interprets the statutory principle in 4.108 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**.

[69] ICO guide to the UK GDPR

[70] ICO Age Appropriate Design Code (The Children's Code)

4.111 This is a priority for VSP providers specialising in pornographic material, VSP providers with services on which there is a significant risk of under-18s encountering pornographic material, and/or VSP providers that allow pornographic material in their terms of service. It is for VSP providers to consider these factors and decide whether robust access controls need to be applied either to the whole platform or a part of it.

4.112 Should Ofcom be required to make an assessment about whether a platform requires such measures, some of the indicators that we might consider in making this assessment include:

- How much pornography is on the platform. This could relate to the absolute number of videos; the ratio of pornographic to non-pornographic content; or the relative number of sub-sections of the site dedicated to pornography.
- The significance of pornography to the service. Where the site allows for users to subscribe to the accounts of content creators, this might relate to how many of those accounts post pornography, or the number of subscribers pornographic accounts have.
- The way the service is positioned in the market. This could be how it brands its own offering, or how the platform is viewed by users.
- Third-party insights which indicate the service specialises in pornography, or that there is a high risk of under-18s being able to access pornographic material on the platform.

4.113 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.[71]

4.114 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 restricted material of a pornographic nature, similar to the approach to R18 or R18-like material.[72]

4.115 When a VSP provider 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*.[73]

4.116 Ofcom takes very seriously the potential risks to under-18s who may be able to access adult platforms and upload content that might be child sexual abuse material. We expect VSP providers that host pornographic material to prioritise the safety and welfare of under-18s and use their access control measures to prevent underage users from accessing the service and uploading material, the inclusion of which could be a criminal offence in the UK.[74]

[71] See BBFC's definition of Sex works at 18.

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

[73] See Protecting under-18s from material unsuitable for classification for a further type of restricted material for which we would expect a robust age verification system to be in place.

[74] Indecent photographs of children – s1, Protection of Children Act 1978. See also section 3.24 of the guidance – 'Material the inclusion of which would be a criminal offence'.

4.117 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.[75]

4.118 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 methods.

4.119 Ofcom would not consider the following to be appropriate forms of age verification 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 or otherwise easily known information such as name, address, and date of birth to verify the age of a user. This does not include electoral roll information, which is a valid data source for age verification.

Protecting under-18s from material unsuitable for classification

4.120 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.[76]

4.121 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.

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

4.122 We recognise that there is a very broad range of 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, based on a literature review of existing evidence. VSP providers may also wish to refer to the BBFC's age-based classification guidelines, which provide helpful

[75] Age Verification Providers Association is the industry trade body for UK age verification providers. Its members are developing a range of solutions that a VSP provider might consider implementing.

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

information about the types of material that might be unsuitable for under-18s, by different age groups.[77]

4.123 VSP providers need 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 and create an age-appropriate experience.

Age assurance and age-appropriateness for under-18s

4.124 In order for VSP providers 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.

4.125 Whilst all platforms should be aiming for an age-appropriate experience for their users, we recognise that there are limits to how far VSP providers 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 VSP providers in being able to protect the youngest and most vulnerable under-18s.

4.126 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 Age Appropriate Design Code, as well as the ICO's more general data protection requirements.[78]

Considerations for effective age assurance

4.127 VSP providers may consider the following factors when establishing and operating age assurance systems:

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) VSP providers should consider how reliable and accurate any age assurance method is and what level of confidence it provides, in relation to the risk. For example, if a solution is not able to accurately distinguish between an adult and a child on a service

[77] BBFC Classification Guidelines

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

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 users can provide false information to easily bypass age assurance measures, e.g. self-declaring to be over the age required by a platforms' terms of service. VSP providers 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) VSP providers should consider how different tools such as reporting of underage accounts, 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 adhering to any technical standards in this area as they develop, such as PAS 1296, the Government's Digital Identity and Attributes Trust Framework[79] or any other relevant standards.[80]

g) When considering age assurance solutions, VSP providers 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, systems based on facial biometrics may carry the risk of algorithmic bias against people from non-white ethnicities, or the age of children with special educational needs may not be easily verifiable through behavioural or profiling methods.

Parental control systems

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

4.128 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

[79] 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.

[80] 5Rights Foundation's report on age assurance, 'But how do they know it is a child?' lays down a comprehensive set of standards for age assurance that VSP providers may consider adopting as good practice.

to under-18s should strongly consider having some form of parental control feature to support their overall protection measures for under-18s.

4.129 There is 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.

4.130 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.

4.131 VSPs may include features that allow parents or guardians to verify the age of child users, such as account verification. Although account verification may not be sufficient as a single means of age assurance on higher risk services, due to some of the limitations with parental control systems set out below, providers offering services with particular appeal to younger users should consider the value of parental consent tools to support their overall child safety approach.

4.132 The limitations to relying on account confirmation from a parent or guardian as a means of verifying a child's age include taking into account the need for active engagement and awareness from parents and guardians, the existence of a trust-based relationship between parent/carer and child, and the risk of inaccurate age assurance if a responsible adult willingly enters a false age for the child and exposes them to age-inappropriate material.

4.133 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.[81]

4.134 Parental control systems can allow guardians to set boundaries for their children online, as they might offline. In principle, they can be used as a way for parents to feel comfortable that their child is using a VSP within safe parameters, rather than being an instrument for monitoring and control. As such, these tools can work most effectively where there is a trust-based dynamic between the parent or carer and the under-18 user. Where there is less trust in the parent/carer – child relationship and/or if children have access to tools to circumvent parental checks (e.g. VPNs, changing login details), parental controls will be less reliable and successful.

4.135 Parents and guardians need to know what parental controls systems VSPs offer and understand how best to use them to support a child's online experience. VSPs could use guides and other media literacy tools alongside parental controls to support this awareness and understanding. As we suggest at 4.173, VSPs may consider best practice from, and

[81] 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.

partnerships with, organisations in the provision of media literacy tools. For example, the ICO and Internet Matters both provide guidance to help parents and guardians to understand Parental Controls.[82]

4.136 VSP providers should be mindful that not all children in the UK have parents or guardians that are able to make use of parental controls to protect their children online. For instance, parents with less familiarity with online services and tools may find it difficult to apply parental controls effectively. Therefore, parental controls should be deployed by VSPs as part of a broader set of measures, rather than being the single available measure for protecting under- 18s.

4.137 VSP providers should also consider the kind of information that is shared with parents, for example, parental controls should not create a situation where a vulnerable child in an abusive household cannot search for abuse helplines without their parents also seeing this.

4.138 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.

4.139 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.

4.140 Adults responsible for under-18s should be given only as much control as is necessary to protect them from harmful material. To consider the best interests of a child, VSPs can design parental controls to ensure under-18s are able to enjoy age-appropriate content and activity on a service without undue parental interference.

4.141 Autonomy online becomes more important as teenagers get older. VSPs should consider the best interests of a child when designing parental controls and how the balance between protection and autonomy might be different as children get older.

4.142 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.

[82] ICO Parental Controls standard; Parental Controls & Privacy Settings Guides - Internet Matters)

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.*

4.143 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 states that 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.

4.144 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.

4.145 This complaints process should be available to all users of a VSP.[83] In the context of this measure, this could include a parent or guardian who 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

4.146 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 explaining: the information required from a complainant to make a complaint; the likely timeframe of the complaint process; potential outcomes of a complaint; and what communication the complainant can expect from the VSP provider throughout the process.

4.147 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.

4.148 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.

[83] Ofcom deems a user as anyone able to access, view or upload content on a platform, not just those who have an account. In the context of certain protection measures, a user could also be a parent or guardian.

4.149 We recommend that data regarding the number of complaints and their outcomes is 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 can support Ofcom's monitoring of the VSP Regime (see Section 8).

4.150 Where a provider has implemented a complaints process, the impartial dispute resolution procedure (as set out from 5.1) and how to access it should be communicated to the complainant at the conclusion of the process.

Easy to use

4.151 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 discourage users from making a complaint. 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).

4.152 VSP providers should consider the platform's user profile when designing a complaints procedure (see Section 6). We expect VSP providers to take an inclusive approach that takes account of the needs of vulnerable users. As explained above in paragraph 4.10, 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.

4.153 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

4.154 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 Section 5).

4.155 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

4.156 It is important to note that the complaints process explained above is distinct from the dispute resolution procedure referred to in Section 5. VSPs have the discretion to require that users exhaust any complaints processes prior to engaging with the dispute resolution process.

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*

4.157 Ofcom defines media literacy as “the ability to use, understand and create media and communications in a variety of contexts”.[84] Media literacy can help to protect users of VSPs from harmful material, and reduce the impact of harmful material, by fostering skills in safer use, critical understanding, and responsible creation, and enabling users to better interpret, curate, and respond to their experiences while using a VSP.

4.158 VSP providers should consider three broad areas when designing and implementing this protection measure:

a) What specific tools and information are needed to improve users’ media literacy based on the nature of the service, the types of users on the service and their specific needs, and the types and prevalence of material that could be harmful on the service;

b) How to raise awareness of the tools and information provided; and

c) How to understand the effectiveness of the tools and information provided on an ongoing basis.

4.159 VSP providers should also consider the five principles set out at paragraphs 4.7 to 4.12 when designing and implementing this protection measure:

- **Effective:** tools and information to improve media literacy should practically help users to protect themselves and others. As such VSP providers should consider the design, delivery and content contained in the tools and information to help achieve the impact they are seeking to have on the skills, knowledge, and behaviours of their users (see understanding effectiveness below at paragraphs 4.163 to 4.164).
- **Easy to use:** tools and information to improve media literacy should be easy to locate, understand and engage with, taking into account the needs of the users that they are aimed at and, if relevant, the device(s) that are likely to be used to access them. Providers might consider when would be most timely to raise awareness of tools and information to increase engagement. VSP providers should have regard to the same considerations as set out under the terms and conditions protection measures in

[84] Information about Ofcom’s media literacy activities

paragraphs 4.28 to 4.40 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 provider plays in protecting its users from harmful material.
- **Evolving:** tools and information to improve media literacy should be regularly reviewed and amended to reflect changing attitudes and behaviours, changes to the service, and new evidence on what is effective as it becomes available.

4.160 Media literacy is an important consideration in the design and implementation of other protection measures, so VSP providers should take a holistic approach to media literacy on their service beyond the provision of specific tools and information. This might, for example, include understanding the media literacy levels of their users when assessing risk and designing and implementing other protection measures.

Improving users' media literacy

4.161 VSP providers should consider the specific information and tools needed to improve users' media literacy based on the nature of their service, the types of users on their service and their needs, and the types of material and prevalence of material that could be harmful on the service. This could include, but is not limited to, providing tools and information that help users develop and exercise knowledge and skills in the three areas outlined below.

a) Using the VSP in a safe and secure manner. This might include:

i) Understanding the safety features available on the VSP, how to use them, and the benefits of doing so, including complaints processes, dispute resolution procedures, parental controls and reporting and flagging mechanisms.

ii) Understanding resources provided by the VSP, or by external organisations, to support and protect users if they are seeking, have been exposed to, have created, or have been impacted by harmful material.

iii) Awareness of the potential benefits and possible risks of using the VSP, including encountering the harms outlined in Section 3 above, and advice on safe behaviour.

b) Critically understanding the VSP and the content available on it. This might include:

i) Understanding how content is delivered to them and, if relevant, how content recommendations are informed and made.

ii) Differentiating between the different types of content available on the VSP (including user-generated content, professionally-generated content and advertising) and, where appropriate, the sources of that content.

iii) Understanding how the VSP is funded and, if relevant, how the user contributes to that funding model (for example, if user data is used to inform targeted advertising).

c) Creating and uploading content responsibly. This might include:

   i) Understanding what content is and is not acceptable on the VSP.

   ii) Understanding the consequences of uploading unacceptable content.

Raising awareness

4.162 VSP providers should consider how to raise awareness of the tools and information they provide to improve media literacy as part of the overall design of the user experience and user journey. We recognise that the specific points at which it might be most effective to raise awareness need to be tested and iterated, but VSP providers should consider how they can raise awareness:

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

   VSP providers should consider how they can best make users aware of the availability of tools and information to improve media literacy at an early point in the user journey. Relatedly, we strongly encourage VSP providers to 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.**

   VSP providers should consider how awareness of tools and information to improve media literacy can be raised at regular intervals throughout the user journey. They should consider whether regular check-ins or prompts are appropriate, particularly if a user has changed default settings, or whether permanent signposting of tools and information is appropriate.

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

   VSP providers should consider how they can raise awareness of resources that can help users respond to harmful material close to the point of engagement with it. This could include signposting users to helpful resources if they search for terms likely to generate harmful material, placing warning labels on harmful material to advise users of its nature and to direct them to helpful resources, signposting users who post or flag harmful material to helpful resources, and providing support to users who viewed harmful material that was subsequently removed.

Understanding effectiveness

4.163 VSP providers should 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 measuring awareness of and engagement with the tools and information, and whether engagement resulted in an improvement in media literacy. These findings can then be used to improve the tools and information provided to make them more effective.

4.164 We recommend that VSP providers also consider best practice from and partnerships with other organisations to ensure they are providing tools and information to improve media literacy in the most effective way. One way to learn from and share learnings with others is through media literacy networks that aim to increase collaboration, information-sharing, and debate among relevant stakeholders to improve media literacy.[85]

Designing effective protection measures - understanding user behaviour

4.165 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 VSP providers to assess whether their users can engage with all measures in a way that protects them from harm.

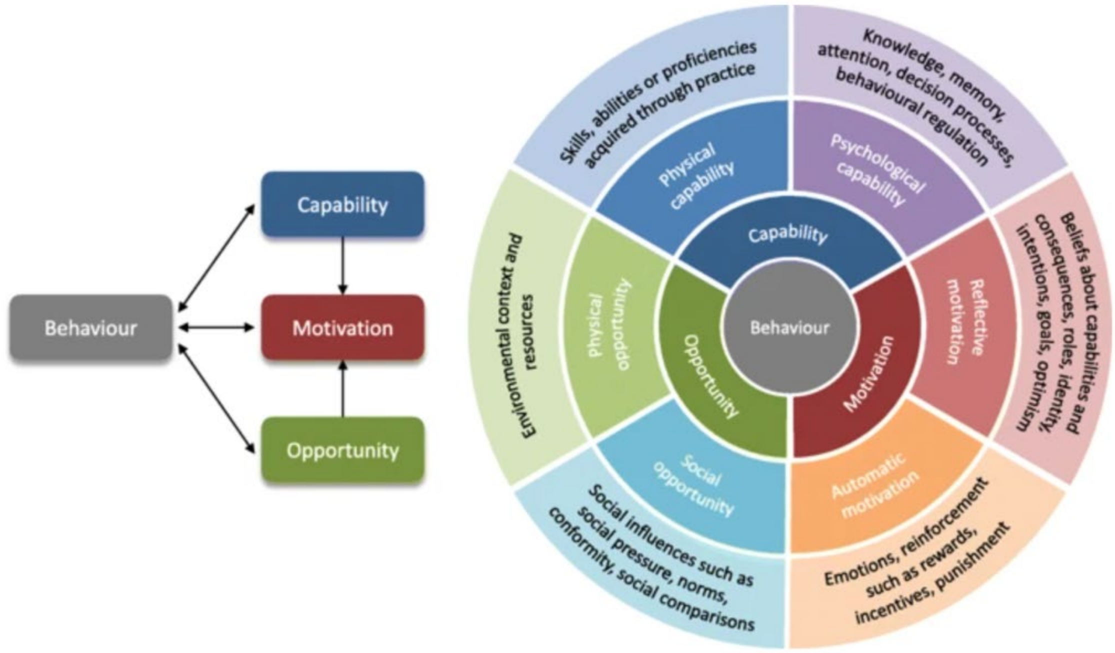
4.166 One way in which VSP providers can achieve this is by systematically analysing whether there may be behavioural factors that could limit the effectiveness of their measures. For example, as mentioned above, users may lack awareness of the measures, or the measures themselves may be designed in such a way that makes engagement difficult. Different groups of users will exhibit different behavioural factors that could limit their ability to engage with protection measures; in particular, VSP providers should be aware of the evolving capabilities of children and their developmental needs meaning that, for example, tools and information to improve their media literacy should be pitched at an appropriate level.

4.167 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.[86] This is illustrated in Figure 4.2 below.

[85] Ofcom have a media literacy network called the Making Sense Of Media Network.

[86] 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.

Figure 4.1: The COM-B model


Capability
Behaviour
Motivation
Opportunity
Skills, abilities or proficiencies acquired through practice
Knowledge, memory, attention, decision processes, behavioural regulation
Physical capability
Psychological capability
Capability
Environmental context and resources
Physical opportunity
Opportunity
Behaviour
Motivation
Reflective motivation
Beliefs about capabilities and consequences, roles, identity, intentions, goals, optimism
Social opportunity
Automatic motivation
Social influences such as social pressure, norms, conformity, social comparisons
Emotions, reinforcement such as rewards, incentives, punishment

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

4.168 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.

4.169 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.

5. Dispute resolution

5.1 The VSP Regime includes a requirement for all providers to implement a dispute resolution procedure, regardless of the size or nature of the platform.[87] This is a separate requirement to the requirement to take and implement appropriate measures to protect users from harmful material. Dispute resolution procedures must be impartial and must allow users to challenge a VSP's implementation of a measure, or a decision to take, or not to take, a measure.

> *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.*

5.2 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).

5.3 No set approach is outlined in the legislation to meet this requirement, providing the procedure: is impartial; is out-of-court; allows for the resolution of user disputes relating to a VSP provider's implementation of any measure or a decision to take, or not to take, any such measure; and does not affect a user's ability to bring a claim in civil proceedings.

5.4 VSP providers have the discretion to decide their approach to the dispute resolution procedure provided it meets the above requirements of the legislation. Ofcom acknowledges that this is an untested area in which existing models by VSP providers may be limited.

5.5 In addition to meeting the statutory criteria listed above, Ofcom recommends that dispute resolution procedures are transparent, easy to use and fair. By this we mean that they should not unduly discriminate between users, introduce bias or result in inconsistent application.

5.6 The dispute resolution procedure is an important requirement in the VSP Regime to ensure the legitimate interests and rights of users are taken into account and balanced against the importance of protection from serious harm.

5.7 Decisions about the implementation of a measure could include a decision concerning the way in which a measure is applied. For example, the application of terms and conditions might result in the removal of a user's content or suspension of a user's account. A decision to take or not to take a particular measure could include a VSP not having parental control mechanisms or age verification systems in place.

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

5.8 In designing and providing a dispute resolution procedure, it may be helpful for VSP providers to consider: how users typically interact with the VSP; the protective measures that are already in place and the sophistication of those measures (for example, whether a robust complaints process is in place to alleviate user concerns); the size of the VSP; the volume of complaints received and decisions taken by the VSP following the consideration of a complaint; the rights and legitimate interests of users who may bring a dispute; and whether an appeals process is already in place and whether this can be modified to provide a dispute resolution procedure that meets the statutory criteria.

5.9 VSP providers may wish to set out criteria that a user should meet in order to submit a dispute and for it to be taken forward. This could include, for example:

a) A requirement that the user must first have exhausted the complaints process (except in cases where the subject of a dispute is not covered by the VSP’s complaints procedure).

b) A requirement that the user must submit the dispute within a reasonable time period after the complaint has been closed by the VSP or, in the case of an issue not covered under the complaints process, the issue has been determined. In setting any timeframe, providers should consider fairness and accessibility.

c) A requirement that the user must include certain specified information, such as the identity of the user or details of the reason for the dispute.

5.10 Depending on the approach a VSP provider takes to ensuring impartiality (see below) the mechanism for resolving a claim which is upheld may vary. For instance, the VSP provider could be required through the dispute resolution procedure to take certain specified steps to resolve the matter, or it might be that these could be for the VSP provider itself to determine where, for example, an external dispute resolution procedure is used.

5.11 Potential ways in which an upheld dispute is resolved could include, for example: issuing an apology, correction or statement; reinstating content; removing sanctions against users; changes to / reviews of processes or policies; introducing or removing particular measures; or some other form of user redress, appropriate to resolve the issue in question.

5.12 The VSP Framework sets out that the use of a VSP provider’s dispute resolution procedure must not affect the ability of a person using the service to bring a claim in civil proceedings.

Ensuring dispute resolution procedures are impartial

5.13 In order to be impartial, platforms must be able to demonstrate procedural separation between any complaint and/or reporting processes that they have in place and their dispute resolution procedure.

5.14 Ofcom considers the following examples to be potential approaches VSP providers could consider taking in relation to ensuring impartiality:

i) an external, fully independent decision-making body or person. This may take, for example, the form of employing the services of an appropriately qualified third party or industry body;

ii) an external independent body or board. This may be established by the provider for the sole purpose of being the dispute resolution procedure for the platform; or,

iii) an internal, procedurally separate person or team.

5.15 Ofcom is aware that third party solutions are an immature but developing sector in the UK and welcomes innovation in this area to expand the range of online dispute resolution options available. 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 and in accordance with the requirements set out in the VSP Framework. At a minimum, Ofcom considers that impartiality requires a separate decision-making process for the resolution of disputes to the process under which the original complaint was handled.

5.16 The VSP Framework is expected to be superseded by future Online Safety legislation (see paragraphs 2.23 and 2.24). The Government has signalled that while companies will be expected to have redress mechanisms, it does not intend to include an independent resolution mechanism in its legislation.[88] In setting out our guidance on the requirement under the VSP Framework to provide for an impartial out-of-court dispute resolution procedure, we are mindful of this and the potential impact on VSPs of differing sizes. We will work with VSP providers to understand and overcome these challenges, including practical timelines for implementation.

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

5.17 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 raise a dispute;

b) information about what will happen once a dispute 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 dispute was considered, and any actions to be taken by the VSP; and

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

[88] Online Harms White Paper response, paragraph 4.43 “Establishing an independent mechanism for resolving disputes would not align with our overarching objective to ensure companies take more responsibility for their users’ safety, and to improve users’ trust in their processes. It could disincentivise cultural change within companies, and encourage companies to ‘offload’ difficult content decisions externally.”

5.18 Dispute resolution procedures should be easy for all users to find and engage with, including provisions for vulnerable users (see 4.10). VSP providers should consider the platform's user profile when designing a dispute resolution procedure (see Section 6).

5.19 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.

5.20 VSP providers maintaining oversight of this procedure can help them to quality assure the enforcement of terms and conditions, mitigate against unwarranted takedown of content, provide transparency and accountability around user safety decisions, and ensure decisions taken reflect the particular needs of users.

5.21 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, as well as any information it has about the user uploading the material.

5.22 Similar consideration should be given to disputes concerning material on a VSP uploaded by a UK-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.

6. Determining which measures are appropriate

6.1 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.[89]

6.2 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.[90]

6.3 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. Ofcom will continue to work with providers and stakeholders to ensure the objectives of the VSP Framework are met.

6.4 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.

6.5 In Section 6 we strongly encourage all VSPs to conduct some form of risk management process and suggest a framework for this. Providers may find that putting in place a process for assessing and managing risks, and collecting information to inform their decisions, is helpful when assessing how they should implement protection measures.

[89] 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.

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

When providing information about the practicable and proportionate criteria below, we do so with these risk management processes in mind.

6.6 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.

Size of the platform

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

6.8 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

6.9 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.

6.10 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).

6.11 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.

6.12 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

6.13 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.

6.14 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 will not be justification for not sufficiently protecting users from harmful material.

Nature of the service

6.15 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

6.16 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.

6.17 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.

6.18 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 7 – Embedding a safety-first approach).

Business and operating models

6.19 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).[91]

[91] “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.

6.20 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.

6.21 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 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

6.22 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.

6.23 VSP providers should have regard to Section 3 where we discuss 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.

6.24 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 discussed in Section 3.

6.25 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.[92]

The characteristics of the category of persons to be protected

6.26 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.[93] Platforms will be better able to tailor their protection measures to the relevant audience using this information.

6.27 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). Providers should implement measures with regard to the best interests of users under the age of 18, where applicable.

[92] See section 368Z1 (5) of the Act

[93] Ofcom deems a user as anyone able to access, view or upload content on a platform, not just those who have an account.

6.28 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 4.10).

6.29 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.[94]

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

6.30 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.

6.31 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 Sections 4 and 5). This is important to ensure that users are not unreasonably impacted by unwarranted takedown of content and that systems are not overly vulnerable to misuse and abuse.

6.32 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 and are regulated by the ICO. These rights and 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, the processing should not be re-used for commercial purpose and must comply with the UK GDPR and the UK Data Protection Act 2018.

Interests of the service provider

6.33 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

[94] ICO Children's Code Hub

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.

6.34 There may be situations where a provider has decided not to take a protection measure for reasons of cost or commercial impact, but that Ofcom believes would be appropriate to protect users from harm on the service. Careful consideration will be given to these situations. As noted above, a lack of resources or an unwillingness to invest in new measures to protect users, are not necessarily justification for not taking a particular measure. 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

6.35 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 5 we suggest ensuring that robust dispute resolution processes are in place which give careful consideration to this content.

7. Additional steps to protect users

7.1 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.

7.2 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.

7.3 We also strongly encourage providers to put in place a process for assessing and managing risk 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.

7.4 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

7.5 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.

7.6 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.

7.7 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 is/are specifically responsible for user safety.

7.8 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; understanding the effectiveness of protection measures; 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 the most egregious or illegal harms that require immediate action.

7.9 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.

7.10 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

7.11 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

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

7.13 **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).

7.14 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.

7.15 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),** or **the Internet Watch Foundation (IWF)**.

Charities, NGOs and harms experts

7.16 We encourage VSP 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. Platforms could also point users to resources created by these experts, including users who have been found to have uploaded harmful material themselves.[95]

7.17 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.

7.18 Providers may consider seeking the views of specialist charities and NGOs when developing their policies and practices, including the **NSPCC**,[96] **Holocaust Educational Trust,**[97] **Tell MAMA**[98] and **Beat.**[99]

7.19 For criminal content such as terrorism and CSAM, we suggest all VSP providers seek specialist advice from organisations dedicated to these harms and we encourage those with a higher risk profile for this type of content to explore more formal partnerships or memberships.

7.20 For advice and support around terrorism, providers may wish to consult with specialist organisations, such as **Tech Against Terrorism** who work with the tech industry to provide the tools needed to tackle terrorist activity on their platforms.

7.21 For CSAM, VSP providers may wish to consult with specialist organisations such as the **IWF**, **NCMEC**, and the **NSPCC,** who are able to offer advice and support on reducing the risk of CSAM on platforms, as well as wider issues relating to child abuse, such as online grooming.[100]

7.22 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. Industry bodies such as the **Age Verification Providers Association**, the **Online Safety Tech Industry Association (OSTIA)** or the **Safety Tech Innovation Network** may be useful sources of advice and information.

7.23 Platforms could consider using trusted flaggers to identify harmful content in a more robust way. It is for a VSP provider to decide what the purpose of a trusted flagger is and who trusted flaggers are, but Ofcom is aware of several organisations operating a service which helps users to report harmful content, as well as providing advice and mediation support. Example organisations include **SWGfL**, **CST** and **Tell MAMA**.

7.24 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

[95] This is because we recognise those uploading harmful material may be in need of support themselves. However, such support should be balanced with the need to take effective action in response to violations of terms and conditions.
[96] The National Society for the Prevention of Cruelty to Children
[97] The Holocaust Educational Trust
[98] Tell MAMA supports victims of anti-Muslim hate and measures and monitors anti-Muslim incidents.
[99] Beat is a UK eating disorder charity.
[100] Though direct messaging is not in scope of the VSP Regime, we are aware that VSPs may be used by offenders to groom children for the purposes of generating new CSAM.

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.

7.25 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

7.26 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 plays a prominent role in the draft Online Safety Bill[101] and is also part of the European Commission's Proposal for a Regulation on a Single Market For Digital Services (Digital Services Act).[102] The VSP Regime is an important first step for Ofcom and industry in online regulation. Following our strong recommendation for VSP providers to implement a risk management framework would provide an opportunity for us all to build expertise, test and iterate to achieve good outcomes for users.

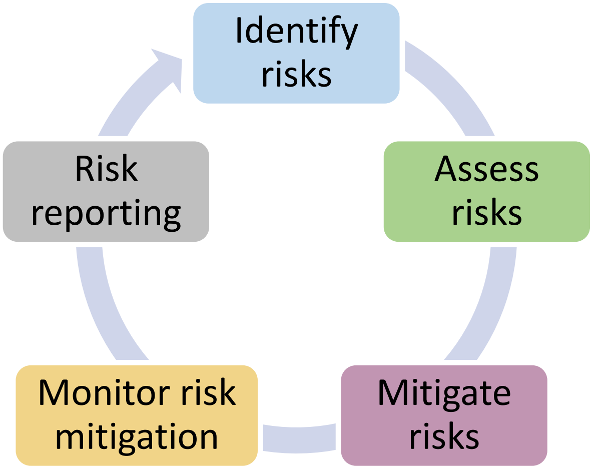
7.27 Below we provide guidance on how VSP providers might conduct this process. This is presented as a step-by-step guide to support providers, but we recognise that individual platforms might approach risk in different ways, dependent on their own particular circumstances.

Risk management process

7.28 The basic steps of a risk management process are **identifying**, **assessing**, planning for and **mitigating** any risks, **monitoring** to ensure that risks are appropriately mitigated. Risk management activities should also be reported to senior decision-makers in the organisation to ensure adequate oversight and governance of the process.

[101] Draft Online Safety Bill, Presented to Parliament by the Minister of State for Digital and Culture by Command of Her Majesty May 2021

[102] The Digital Services Act: ensuring a safe and accountable online environment



7.29 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 recording 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.

7.30 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 demonstrably satisfy themselves that they have adequately considered the risks of harmful material on their platform in determining whether the measures they take or have taken are appropriate, taking into account the practicable and proportionate criteria, and protect users.

7.31 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. Our intention is to understand providers' existing risk and safety management processes, including those they might have in place to meet other regulatory requirements. We will then discuss with providers (where appropriate) how these can be developed to better protect users under the VSP Regime. This 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

7.32 We encourage providers to carefully consider the existing 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.

7.33 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).[103]
- 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)

7.34 In addition, evidence of risk could come from historical experience of harmful material on the platform, for example through engagement with users via complaints and the reporting and flagging mechanisms providers may have in place.

7.35 We also encourage providers to be alert to the possibility of emerging risks. An emerging risk is a new or unforeseen risk that has not yet been fully understood.[104] We do not expect platforms to have perfect foresight or to be able to anticipate and identify all emerging risks. We do consider though, that it is reasonable for a platform to make attempts to identify and, where it is proportionate to do so, be alert to emerging risks so that they can react quickly, should they need to.

7.36 The identification of emerging risks could be informed by, for example, being aware of the risks that have arisen on similar services, engagement with industry bodies, Ofcom and cross-sector initiatives or workshops, and ongoing engagement with expert bodies such as charities, NGOs and harms experts (see paragraphs 7.16 to 7.25).

7.37 Identification of risks should be performed regularly and systematically to ensure that both existing and emerging risks are considered. We encourage VSP providers to collect relevant information on this process so that this can be reported to senior decision-makers in the organisation to facilitate robust internal conversations that can drive forward a culture of safety.

7.38 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 also be useful as part of compliance monitoring.

[103] 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".

[104] There are a number of ways that emerging risk has been defined. Lloyds of London defines this as "an issue that is perceived to be potentially significant but which may not be fully understood". Swiss Re defines emerging risks as "newly developing or changing risks that are difficult to quantify". The International Risk Governance Council (IRGC) defines an emerging risk as "a new risk, or a familiar risk in a new or unfamiliar context". The International Actuarial Association defines an emerging risk as "a risk which may develop or which may already exist that is difficult to quantify or may have a high loss potential". The Institute of risk management (IRM) defines an emerging risk as a "risk that is evolving in areas and ways where the body of available knowledge is weak". Common features of these definitions are that an emerging risk is a new or evolving unforeseen risk whose significance is uncertain and not fully understood.

Assess the risks of harmful material on the platform

7.39 Having identified the relevant existing and emerging risks to users encountering harmful material on the platform, it is important that VSP providers consider the likelihood and impact of those risks materialising.

7.40 Some of the practicable and proportionate criteria set out in Section 6 may provide useful context 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 in the risk profile because they are a clear category of persons to be protected under the VSP Framework (see Section 3).

7.41 Understanding the risk profile of their service will help a VSP provider 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 encourage VSP providers to document their risk profile so that this can be reported to senior decision-makers in the organisation, ensuring they have appropriate oversight of the assessment.

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

7.42 Once a VSP provider understands their service's unique risk profile (based on existing and emerging risks), 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?

7.43 We would expect a VSP provider to specifically consider all measures listed in Section 4 that it has implemented, as well as any other relevant protection measures.[105]

7.44 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 their users against harmful material. In other words, in

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

this step, the VSP provider should consider whether the measures, in the round, need to be improved. 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, noting that some rights and legitimate interests of users may be impacted by particular options (see paragraphs 6.30 to 6.35); and

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

7.45 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.

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

7.46 Good risk management is an ongoing process. The risks of harmful material on a VSP are likely to be constantly changing as a result of both internal and external factors. For example, the evolving nature of technology used by a provider and the increasing sophistication of users to circumvent controls means that VSP providers should regularly monitor the effectiveness of their protection measures and evolve, improve or change them to continue to protect their users.

7.47 For VSP providers to satisfy themselves that they are meeting their obligations, VSP providers are strongly encouraged to regularly collect information about the impact of their protection measures and review that against their platform-specific needs, to manage risk and protect users. For example, if a VSP provider has decided to implement a media literacy protection measure because it is deemed important to protect users, then it is reasonable to expect them to consider whether this results in an improvement to users' media literacy and so therefore does protect users. Further detail on the type of information that might be relevant can be found below. If existing measures are not effective at mitigating risks of harm, providers should consider whether further measures, or an extension of existing measures, are needed.

Report risk management activities to senior decision-makers to ensure appropriate oversight and governance of activities

7.48 Risk reporting is the means for communicating the value of risk management activities to senior decision-makers within the organisation. It provides the mechanism for robust discussion and debate on how existing and emerging risks are being assessed and managed internally. In particular, it can be used to provide demonstrable assurance to the senior decision-makers in smaller firms, the Board in larger firms and, in certain circumstances to the regulator, that all relevant risks are being assessed and managed adequately. It might also be useful for platforms to publicise their risk management activities to their users and other stakeholders. This occurs in some industries because greater transparency is likely to result in wider confidence in an organisation's approach to managing risk. For VSP providers, transparency about risk management may build trust in the service and confidence that appropriate protections are in place.

7.49 This step can help ensure risk management is embedded into governance processes and that there is effective oversight implemented across all stages of the risk management framework; it also ensures Board accountability in larger firms and senior decision-maker accountability in smaller firms. For example, when risk reporting occurs alongside changes of the provider's risk profile, management and senior decision-makers of the organisation can make informed decisions as risks change, to ensure good outcomes for both users as well as the business. Further, when risk reporting to senior decision-makers comprises of signposting existing and emerging risks and assessing whether they have been effectively mitigated, this can improve future risk management. This should therefore allow for proactive risk management as organisations identify and escalate issues either as they arise, or before they are realised.

Assessing effectiveness of protection measures

7.50 Whether or not a platform puts in place a risk management process, and follows the framework above, Ofcom considers that assessing 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 and support our aim of raising standards in user protections.

7.51 We strongly encourage VSP providers to collect 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. In the case of tools and information to improve media literacy, platforms should consider collecting information about how effective they are at improving media literacy. Such information is likely to support a VSP provider's risk management system and related decisions about taking further measures or strengthening existing ones.

7.52 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 warnings and bans given to users, or groups of users. This could also include an assessment of how many warnings or temporary bans turn into permanent bans, as an indication of the effectiveness of those initial actions

e) The number of content-related decisions appealed, as well as the outcome of other disputes made under the dispute resolution procedure

f) User awareness of and engagement with protection measures

7.53 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.[106]

7.54 We recognise that the relevant context must be considered along with numerical indicators and we encourage platforms to consider qualitative indicators when evaluating the effectiveness of protection. 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. Equally, changes in the number of videos reported on a platform could be due to the content of the videos themselves, the prevalence of the videos, a change to the platform's processes, or increased engagement with reporting and flagging tools by users.

7.55 Along with qualitative indicators that relate to the platform itself, external events and changes in wider societal trends might have an impact on the number of reports and removals (e.g. terror incidents leading to rise in reports of incitement to hatred). We encourage platforms to be aware of the impact of such events on the information they collect. These events and trends can often be localised and it is good practice for platforms to attempt to collect data split by region or territory.

7.56 Many in-scope platforms publish information about the material reported to them and the subsequent actions taken. Examples include **Snapchat**, **TikTok** and **Twitch**.[107] These reports demonstrate that there are a variety of ways to categorise harm and collect data about reports and subsequent actions. Ofcom's Annual VSP Report will aim to bring together this varied information and present it as part of a coherent picture of how platforms are achieving protections for users. We discuss our Annual VSP Report in more detail in the next section.

[106] Platforms might want to understand the extent to which under-18s are viewing age-inappropriate content to help assess the effectiveness of measures designed to help protect under-18s from Restricted Material. If doing so, providers should have regard to the ICO's Age Appropriate Design Code.

[107] Other online services not in-scope of the VSP regime also publish transparency reports, including: Facebook and Instagram; reddit; Twitter; and YouTube.

8. Ofcom's approach to monitoring and enforcement

8.1 Ofcom has a duty to take steps to ensure that VSP providers comply with their obligations under the VSP Framework, which include taking appropriate measures to protect users from harmful material.[108] Further, Ofcom has set out four broad aims for the VSP Regime, which are: to raise standards in user protections; to identify and address areas of non-compliance; to increase transparency across the industry; and to prepare for Online Safety.

8.2 One of the ways we will meet our legislative duty and achieve our aims for the VSP Regime, is through monitoring and enforcement. In this section we set out some of the ways we currently expect to monitor compliance, including through information gathering and the use of our complaints webform.

8.3 Ofcom has the power to take enforcement action in a number of scenarios, including where we suspect a provider has failed to take or implement appropriate measures to protect users against harm. In this section we provide some information on how Ofcom will approach enforcement.

8.4 Ofcom also has the power to request information for the purposes of publishing reports. These reports can cover the steps taken by VSP providers to comply with their duties, including the measures taken to protect users. We plan to publish an Annual VSP Report every autumn and we discuss this below.

Monitoring

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

8.6 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.

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

8.8 Through the information we gather, we might also identify areas of potential non-compliance. Where this happens, we will follow the approach set out under the Enforcement section below.

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

Supervisory engagement with platforms

8.9 Each year Ofcom will identify priority areas of focus, each of which is likely to be centred on a particular harm or measure. We will adopt a structured programme of supervisory engagement, tailored to individual platforms. We plan on writing to the VSP providers relevant to each area of focus, setting out our expectations for the year ahead.[109]

8.10 This supervisory engagement will support Ofcom's aim of raising standards in user protections and we will seek to understand the measures providers have implemented to reduce and mitigate the risk of harmful material.

8.11 In addition to the planned engagement for our priority areas, other matters might come to Ofcom's attention that mean we will want to engage with platforms. This might require us to gather information from a VSP provider, formally or informally, in order to better understand a particular issue (see Information Gathering below).

8.12 The purpose of informal engagement with platforms is not to identify compliance concerns so that we can take immediate enforcement action. We encourage VSP providers to engage with Ofcom proactively about their compliance concerns. We want to ensure that providers understand their new obligations and how they can comply with them. Informing us of obstacles that providers face in complying can be helpful in case there is support Ofcom can provide in overcoming these challenges.

8.13 We also encourage platforms to bring potential compliance failures to our attention. Ofcom will generally seek to resolve issues informally with platforms, where appropriate. Even where such information does result in a formal investigation, self-admission, particularly at an early stage, is a key consideration in our enforcement guidelines (see Enforcement below).

8.14 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. Should we raise concerns with a VSP provider, we will give a reasonable time to respond to these and for any appropriate discussions to take place.

Engagement with other stakeholders

8.15 Information about potential compliance concerns, or about where harm is occurring online might come to Ofcom's attention through our own monitoring activities. But we will also be working closely with other key stakeholders including other regulators, charities, NGOs and harms experts to identify issues. We will use discussions with these stakeholders to gather wider expertise and insights to inform our regulatory activities. In some cases we will look to create collaborative working groups to tackle specific harms or to provide further guidance on some of the measures in the VSP Framework.

[109] Our priority areas of focus for the first year of the VSP Regime are set out in Ofcom's VSP Plan and Approach document.

Information gathering

8.16 Ofcom has broad statutory powers to request information from VSP providers for the purposes of fulfilling its functions as the regulator.[110] This includes information needed for monitoring providers' compliance with the VSP Framework.

8.17 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 and, in the case of tools and information to improve media literacy, their effectiveness at improving media literacy).

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

8.18 For (b) we have suggested ways a platform might assess effectiveness above at paragraph 7.50. Guidance on risk management processes is set out at 7.26.

8.19 When issuing information requests Ofcom will clearly set out the purpose for the request, the reasons we consider we need the information requested and we will provide the person being asked for the information the opportunity to make representations concerning the request. We will also ensure any request for information is justifiable, proportionate and fair. Insofar as we may need to publish the information, such as in an enforcement notice or in Ofcom's Annual VSP Report (see below), we will have regard to the need to exclude from publication, so far as practicable, matters which are confidential.[111]

Accessing VSPs

8.20 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.[112]

Complaints

8.21 Ofcom does not have a role in responding to or adjudicating on individual user complaints. Individuals should complain about harmful videos to the VSP provider in the first instance. We want users to be engaged with the reporting, flagging and complaints processes of

[110] 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.

[111] In accordance with sections 36810(5) and 368Z11(2) of the Act

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

VSPs. Where users are not satisfied with the safety processes or complaints functions of a VSP, they can let us know via our website.

8.22 Users can submit complaints to us via our online complaints portal. 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.

8.23 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.

8.24 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

8.25 Where we identify compliance concerns, we will assess the issue and consider whether it is appropriate to take enforcement action to help protect users from harm.

8.26 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

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

8.27 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, including potentially by under-18s;
- 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.

8.28 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 investigating further and considering the use of our 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.

8.29 We would be much more likely to consider taking immediate 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.

8.30 In the event of a compliance concern regarding a provider's use of appropriate measures, 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) How decisions were made about which measures to take, or not to take, and decisions about the way in which any measures have been implemented.

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

8.31 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 issue an enforcement notification requiring the VSP provider to take specified actions, and/or impose a financial penalty. Ultimately, we also have the power to suspend or restrict a service in cases involving the most serious non-compliance.

8.32 We will use our enforcement tools proportionately and will always have regard to rights to freedom of expression and to privacy, particularly where more intrusive regulatory interventions are required, such as directions to remove pieces of content or the suspension or restriction of a service.

8.33 Any enforcement action will be taken in line with our Enforcement Guidelines.[113] Ofcom also has published information on the setting of financial penalties in our Penalty

[113] These Guidelines will be updated to reflect new powers Ofcom has been given. Ofcom will update this guidance with the new Enforcement Guidelines when they are published, following consultation.

Guidelines. The considerations set out in both sets of guidelines will be applied as appropriate to any VSP enforcement notice Ofcom issues.

Annual VSP Report

8.34 Ofcom will publish an annual report about steps taken by VSP providers to comply with their duties, including the measures taken to protect users.

8.35 Ofcom will closely engage with UK-established VSP providers to help determine the relevant data and information we will collect for the purposes of producing the report, in addition to our own and third party research.

8.36 In line with our information gathering powers, we are able to request a range of data or information on, or related to, any of the following:

a) The measures taken by providers to protect users from harmful material and the ways in which such measures are implemented to achieve this.

b) 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.

c) Other steps taken by VSP providers to comply with requirements of the VSP Framework, such as the advertising specific requirements, requirements relating to the transparency of advertising, and the requirement for an impartial dispute resolution procedure.[114]

8.37 Ofcom's annual report will provide commentary on industry progress in implementing measures to protect users, report on progress against our published priorities, and recommend priorities for the following year.

8.38 Reports may also include Ofcom's own evidence gathered from our complaints data, VSP consumer research panel, as well as relevant research by third parties.

[114] See sections 368Z10 and 368Z11 of the Act