

Response to Ofcom

Guidance for VSP providers on measures to protect users from harmful material

June 2021

1. About ISBA

- 1.1. ISBA is the only body in the UK that enables advertisers to understand their industry and shape its future, because it brings together a powerful network of marketers with common interests, empowers decision-making with knowledge and insight and gives single voice to advocacy for the improvement of the industry.
- 1.2. ISBA is a member of the Advertising Association and represents advertisers on the Committee of Advertising Practice and the Broadcast Committee of Advertising Practice, sister organisations of the Advertising Standards Association, which are responsible for writing the Advertising Codes. We are also members of the World Federation of Advertisers. We are able to use our leadership role in such bodies to set and promote high industry standards as well as a robust self-regulatory regime.

2. Overview

- 2.1. We welcome the opportunity to make a short submission to this consultation. As an organisation, ISBA has long supported policymakers' and regulators' ambition to make the UK one of the safest places in the world to be online. It is welcome that, for the first time, VSPs which allow users to upload content will be required to adhere to regulation.
- 2.2. In supporting measures to tackle online harms, we have also supported the principle of proportionality; and that the bigger the platform, the greater its responsibilities to ensure the welfare of its users. We agree that this is an important factor in this case of this guidance, and that the biggest VSPs should make a commensurately bigger effort to prevent, identify, and mitigate harm. We also believe that they have the responsibility to monitor the effectiveness of their efforts, to report on them, and to evaluate and improve.
- 2.3. The larger platforms have more video content being uploaded than it is humanly possible to manage, so they are of course increasingly reliant on automation. It is vital that such systems continue to be robustly developed in order to help manage the continuing growth in the number and volume of uploads. With the amount of exposure to harmful content online – and with the sometimes horrendous consequences of that exposure, from mental ill health to physical harm – we and our members strongly agree that it is imperative that consumers and users are protected.
- 2.4. Whilst the major platforms have introduced a range of initiatives aimed at protecting their users from harmful content, the ability to compare the consistency and effectiveness of these measures is questionable. Further, in the various workstreams which we have taken part in alongside government to tackle child sexual exploitation and terrorist content, it has been clear that issues are prevalent across smaller and emerging platforms, supported by advertising from some irresponsible actors. Setting a regulatory bar and reach that addresses this issue is critical. Up until recently, the regulatory environment covering online content has been fragmented and multi-

territorial, and has been driven by public pressure rather than proactive engagement. It has also lacked robust data. We therefore fully support progress toward new regulation, enforced by Ofcom.

- 2.5. The advertising and marketing industry is already working of its own accord to address the challenge of harmful content on digital media platforms, and its monetization via advertising. The Global Alliance for Responsible Media (GARM) is a cross-industry initiative which has been working to set standards and definitions of harmful content. We highlight more about this work below.
- 2.6. We note that this VSP regime is the first step in a process which will continue through, and be superseded by, the online harms regime enabled by the Online Safety Bill. As this will be the foundation of that future framework, we agree that it is important to set fundamental principles now which can stand us in good stead for the future, and accommodate changes in VSP technology and digital advertising.

3. Response

- 3.1. In this response, we are focusing on answering questions 1-7 and 11, where we feel we can offer a view and where responses are of direct relevance to the interests of our members.

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

- 3.2. We note Section 3 of the draft guidance, which offers definitions of the types of material consumers should be protected from. We also note the split of harmful material into two broad categories – ‘restricted’ and ‘relevant harmful’ material.
- 3.3. The definition of relevant harmful material includes content likely to incite violence or hatred against individuals or groups, based on recognized protected characteristics; and material which would be a criminal offence, such as terrorism, child sexual abuse, racism, and xenophobia. ISBA has a history of partnering with policymakers on behalf of our industry to tackle these issues – not simply from a position of brand safety and reputation, but because we believe in a trusted, responsible advertising and online environment. We therefore welcome the definition of relevant harmful material, especially at a time of increased hate speech and disinformation online.
- 3.4. We note that restricted material is that which has, or would be likely to be given, an R18 certificate (or which has been, or would be likely to be, refused a certificate entirely). We presume that the requirement will be for platforms to map their current definitions to that of the BBFC and classify videos accordingly.
- 3.5. We would raise the question of possible exemptions to what would otherwise be R18 category content. For example, there has been discussion when it comes to the upcoming online harms framework around due consideration for material which is for journalistic purposes, and would ask whether this will be taken into account in the VSP regime. We would also raise the issue of the intent of the material which is uploaded. For example, there may be videos which condemn acts of violence; or videos containing images of self-harm which may be uploaded by health bodies or NGOs who are providing support to vulnerable individuals. An intelligent approach to this content would clearly be welcome.

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

- 3.6. We agree that the implementation of a common framework of definitions of what counts as restricted definitions is essential. We would balance this by suggesting that it is also essential that the VSP should notify the user of such material, and that without both sides of this coin, the guidance will be ineffective.
- 3.7. We further agree that the definitions of harmful content are also key to the regime; and that the recommendations for ensuring that terms and conditions are easy to use, effective, fair, transparent, and are required to evolve, are all sound.
- 3.8. At this stage it may be helpful for us to refer to the definitions of harmful content which have been agreed by GARM in their brand safety and sustainability framework. Marketers, agencies, advertising representatives (including ISBA), and platforms including Facebook, YouTube and Twitter have agreed to adopt this common framework for defining harmful content that is inappropriate for advertising, and to collaborate with a view to monitoring industry efforts to improve in this area.
- 3.9. Four key areas for action were identified by GARM, designed to boost consumer and advertiser safety with agreed individual timelines for each platform to implement across the different areas. These are: the adoption of GARM definitions for harmful content for safe advertising practices; the development of more harmonised reporting on harmful content; a commitment to have independent oversight on brand safety operations, integrations and reporting; and a commitment to develop and deploy tools to better manage advertising adjacency.
- 3.10. Historically, definitions of harmful content varied by platform. GARM's Brand Safety Floor and Suitability Framework offers common definitions to which participants have agreed to adhere. The Safety Floor (Fig. 1) lists content for which industry considers that it is not appropriate for there to be any advertising support (and in the context of this consultation, would seem analogous to Ofcom's 'relevant harmful' material). The Suitability Framework (Fig. 2) lists sensitive content which may be appropriate for advertising, when that advertising is supported by proper controls (perhaps analogous to Ofcom's 'restricted' material).
- 3.11. This initiative by industry builds on the self- and co-regulatory system and solutions which are the hallmark of the United Kingdom's successful and world-leading regulation of advertising content. We hope that this framework is of use to Ofcom as a point of comparison and inspiration for the definition of what counts as relevant harmful and restricted content, and for the nuances which can take place when it comes to the interpretation of the impact of restricted content's being consumed by a user.
- 3.12. It will be noted that GARM have agreed that where content is part of the Brand Safety Floor, it is not appropriate for it to be supported by any advertising content. If there is to be read across to Ofcom's relevant harmful material, then it stands to reason that the same consideration may apply. Where content is classed as relevant harmful, it may be useful to mandate that that content should not be monetised.

CONTENT CATEGORY	BRAND SAFETY FLOOR – Content not appropriate for any advertising support
Adult & Explicit Sexual Content	<ul style="list-style-type: none"> • Illegal sale, distribution, and consumption of child pornography • Explicit or gratuitous depiction of sexual acts, and/or display of genitals, real or animated
Arms & Ammunition	<ul style="list-style-type: none"> • Promotion and advocacy of Sales of illegal arms, rifles, and handguns • Instructive content on how to obtain, make, distribute, or use illegal arms • Glamorization of illegal arms for the purpose of harm to others • Use of illegal arms in unregulated environments
Crime & Harmful acts to individuals and Society, Human Right Violations	<ul style="list-style-type: none"> • Graphic promotion, advocacy, and depiction of willful harm and actual unlawful criminal activity – Explicit violations/demeaning offenses of Human Rights (e.g. human trafficking, slavery, self-harm, animal cruelty etc.), • Harassment or bullying of individuals and groups
Death, Injury or Military Conflict	<ul style="list-style-type: none"> • Promotion, incitement or advocacy of violence, death or injury • Murder or Willful bodily harm to others • Graphic depictions of willful harm to others • Incendiary content provoking, enticing, or evoking military aggression • Live action footage/photos of military actions & genocide or other war crimes
Online piracy	<ul style="list-style-type: none"> • Pirating, Copyright infringement, & Counterfeiting
Hate speech & acts of aggression	<ul style="list-style-type: none"> • Behavior or content that incites hatred, promotes violence, vilifies, or dehumanizes groups or individuals based on race, ethnicity, gender, sexual orientation, gender identity, age, ability, nationality, religion, caste, victims and survivors of violent acts and their kin, immigration status, or serious disease sufferers.
Obscenity and Profanity, including language, gestures, and explicitly gory, graphic or repulsive content intended to shock and disgust	<ul style="list-style-type: none"> • Excessive use of profane language or gestures and other repulsive actions that shock, offend, or insult.
Illegal Drugs/Tobacco/e-cigarettes/Vaping/Alcohol	<ul style="list-style-type: none"> • Promotion or sale of illegal drug use – including abuse of prescription drugs. Federal jurisdiction applies, but allowable where legal local jurisdiction can be effectively managed • Promotion and advocacy of Tobacco and e-cigarette (Vaping) & Alcohol use to minors
Spam or Harmful Content	<ul style="list-style-type: none"> • Malware/Phishing
Terrorism	<ul style="list-style-type: none"> • Promotion and advocacy of graphic terrorist activity involving defamation, physical and/or emotional harm of individuals, communities, and society
Debated Sensitive Social Issue	<ul style="list-style-type: none"> • Insensitive, irresponsible and harmful treatment of debated social issues and related acts that demean a particular group or incite greater conflict;

Fig. 1. GARM Brand Safety Floor

3. Regarding terms and conditions which prohibit relevant harmful material, do you have any comments on Ofcom’s view that effective protection of users is unlikely to be achieved without having this measure in place and it being implemented effectively?

- 3.13. The definitions of harmful content laid out in the consultation document appear to us to be logical. Of course, a balance has to be struck between seeking to apply these rules to a broad set of video providers, offering different services and access to a wide range of content. However, the effective self-attestation or application is perhaps too broad.
- 3.14. Notwithstanding this, we believe that an important question is how such harmful material can be prevented from being uploaded in the first place – preventing harm before it is committed (while of course appreciating that, as the document states, it is impossible to promise to prevent *all* harm). We believe that VSPs should be mandated to remind those uploading content that harmful content is disallowed, and that the definitions of harmful content should be signposted to uploaders.
- 3.15. Section 4.18 of the draft guidance states that terms and conditions “should explain the type of content considered to be restricted material ... and specify that videos containing this material must be brought to the attention of the VSP provider”. We would question whether this responsibility should rather rest in the control of the platform, which should dictate agreed terms about what is and is not allowable content

– and implementing controls asking uploaders to certify what type of material they are submitting.

CONTENT CATEGORY	High Risk	Medium Risk	Low Risk
Adult & Explicit Sexual Content	<ul style="list-style-type: none"> • Suggestive sexual situations requiring adult supervision/approval or warnings • Full or liberal Nudity 	<ul style="list-style-type: none"> • Dramatic depiction of sexual acts or Sexuality issues presented in the context of entertainment • Artistic Nudity 	<ul style="list-style-type: none"> • Educational, Informative, Scientific treatment of sexual subjects or sexual relationships or sexuality
Arms & Ammunition	<ul style="list-style-type: none"> • Glamorization /Gratuitous depiction of illegal sale or possession of Arms • Depictions of sale/use/distribution of illegal arms for inappropriate uses/harmful acts 	<ul style="list-style-type: none"> • Dramatic depiction of weapons use presented in the context of entertainment • Breaking News or Op-Ed coverage of arms and ammunition 	<ul style="list-style-type: none"> • Educational, Informative, Scientific treatment of Arms use, possession or illegal sale • News feature stories on the subject
Crime & Harmful acts to individuals and Society, Human Right Violations	<ul style="list-style-type: none"> • Depictions of criminal/harmful acts or violation of human rights 	<ul style="list-style-type: none"> • Dramatic depiction of criminal activity or human rights violations presented in the context of entertainment • Breaking News or Op-Ed coverage of criminal activity or human rights violations 	<ul style="list-style-type: none"> • Educational, Informative, Scientific treatment of crime or criminal acts or human rights violations • News feature stories on the subject
Death, Injury or Military Conflict	<ul style="list-style-type: none"> • Depiction of death or Injury • Insensitive and irresponsible treatment of military conflict, genocide, war crimes, or harm resulting in Death or Injury • Depictions of military actions that glamorize harmful acts to others or society 	<ul style="list-style-type: none"> • Dramatic depiction of death, injury, or military conflict presented in the context of entertainment • Breaking News or Op-Ed coverage of death, injury or military conflict 	<ul style="list-style-type: none"> • Educational, Informative, Scientific treatment of death or injury, or military conflict • News feature stories on the subject
Online piracy	<ul style="list-style-type: none"> • Glamorization /Gratuitous depiction of Online Piracy 	<ul style="list-style-type: none"> • Dramatic depiction of Online Piracy presented in the context of entertainment • Breaking News or Op-Ed coverage of Online Piracy 	<ul style="list-style-type: none"> • Educational, Informative, Scientific treatment of Online Piracy • News feature stories on the subject
Hate speech & acts of aggression	<ul style="list-style-type: none"> • Depiction or portrayal of hateful, denigrating, or inciting content focused on race, ethnicity, gender, sexual orientation, gender identity, age, ability, nationality, religion, caste, victims and survivors of violent acts and their kin, immigration status or serious disease sufferers, in a non-educational, informational, or scientific context 	<ul style="list-style-type: none"> • Dramatic depiction of hate speech/acts presented in the context of entertainment • Breaking News or Op-Ed coverage of hate speech/acts 	<ul style="list-style-type: none"> • Educational, Informative, Scientific treatment of Hate Speech • News features on the subject
Obscenity and Profanity, including language, gestures, and explicitly gory, graphic or repulsive content intended to shock and disgust	<ul style="list-style-type: none"> • Glamorization /Gratuitous depiction of profanity and obscenity 	<ul style="list-style-type: none"> • Dramatic depiction of profanity and obscenities presented in the context of entertainment by genre • Breaking News or Op-Ed coverage of profanity and obscenities Genre based use of profanity, gestures, and other actions that may be strong, but might be expected as generally accepted language and behavior 	<ul style="list-style-type: none"> • Educational or Informative, treatment of Obscenity or Profanity • News feature stories on the subject
Illegal Drugs/Tobacco/e-cigarettes/Vaping/Alcohol	<ul style="list-style-type: none"> • Glamorization /Gratuitous depictions of illegal drugs/abuse of prescription drugs • Insensitive and irresponsible content/treatment that encourages minors to use tobacco and vaping products & Alcohol 	<ul style="list-style-type: none"> • Dramatic depiction of illegal drug use/prescription abuse, tobacco, vaping or alcohol use presented in the context of entertainment • Breaking News or Op-Ed coverage of illegal drug use/prescription abuse, tobacco, vaping or alcohol use 	<ul style="list-style-type: none"> • Educational, Informative, Scientific treatment of illegal drug use/prescription abuse, tobacco, vaping or alcohol • News feature stories on the subject
Spam or Harmful Content	<ul style="list-style-type: none"> • Glamorization /Gratuitous depiction of Online Piracy 	<ul style="list-style-type: none"> • Dramatic depiction of Spam or Malware presented in the context of entertainment • Breaking News or Op-Ed coverage of Spam or Malware 	<ul style="list-style-type: none"> • Educational, Informative, Scientific treatment of Spam or Malware • News feature stories on the subject

Fig. 2. GARM Suitability Framework

4. Do you have any comments on Ofcom's view that, where providers have terms and conditions requiring uploaders to notify them if a video contains restricted material, additional steps will need to be taken in response to this notification to achieve effective protection of under-18s, such as applying a rating or restricting access?

- 3.16. We agree with the assertion that controls on the uploading of material, married with age-gated controls managing the accessing of such content by the user, are necessary to create a fully effective system.
- 3.17. Answering the challenge set out in the consultation, as well as many other relevant public policy issues, relates to the critical question of age and ensuring both that minors are protected from harmful content, and that those accessing certain content have satisfactorily demonstrated their age. Regulators and policymakers have issued relevant consultations here, including the ICO's on the Age Appropriate Design Code, and the Government's recent National Data Strategy.
- 3.18. When it came to the Age Appropriate Design Code, the compliance options set out were either to apply standards to all users by default, or to offer a robust and effective age-verification mechanism. The recommendation was for the latter, in order to tailor the experience for each age range.
- 3.19. While the ICO was clear that age verification does not solve every issue, the Code inevitably encourages much greater use of it. It was our belief that this was likely to deliver one of two consequences: requiring significantly more data collection from both adults and children, in direct contradiction to the principle of data minimisation; or, potentially, arbitrary age-gating that would restrict children's ability to benefit from the online world.
- 3.20. ISBA and our members have long sought improved enforcement of age limits through improved verification techniques. We continue to believe that further innovation is needed in this area to improve standards, and there are a range of techniques which could be deployed, up to and including formal identification through documentation.
- 3.21. The challenges identified in the National Data Strategy underline the potential for age verification and associated digital identities; the Government noted this in the context of allowing people to verify their identity for the purposes of individual transactions, from applying for benefits to buying a house. It is possible that a trusted, verifiable digital identity could also be a powerful tool for clamping down on online harms such as trolling, or give platforms greater ability to prevent younger age groups from seeing advertising which would be inappropriate for them.
- 3.22. There remains no consensus on a particular solution for this, but when it comes to the additional steps which need to be taken to achieve effective protection of under-18s, we believe that it is important that a system of age verification continues to be explored – in order for platforms to take more meaningful action to protect vulnerable people, properly target advertising, and promote accountability.

5. Do you have any comments on the draft guidance about reporting or flagging mechanisms, including on Ofcom's view that reports and flagging mechanisms are central to protecting users?

3.23. Our general comment is that these systems, which allow users to rate and/or report harmful material, must be designed so that they are intuitive and easy to use. We would encourage Ofcom to provide guidance for VSPs on what rating system they should be using, and that wherever possible this should be consistent across platforms, rather than being left to their discretion.

6. Do you have any comments on the draft guidance about systems for viewers to rate harmful material, or on other tagging or rating mechanisms?

3.24. It is possible that the obligations being set out in the guidance are too loose, and that it is not easy to distinguish between the individual uploading content, and the user downloading it. When it comes to the person uploading, it is our view that at the point at which the upload is made, there should be a clear notice indicating an uploader's obligations under the platform's terms and conditions, including what counts as relevant harmful and restricted material.

3.25. VSPs should also be taking measures to determine, through content verification processes and techniques, what content is actually being uploaded to their platform. As it stands, in our view it is not clear that this is an obligation being levelled on VSPs.

7. Do you have any comments on the draft guidance about age assurance and age verification, including Ofcom's interpretation of the VSP Framework that VSPs containing pornographic material and material unsuitable for classification must have robust age verification in place?

3.26. Please see our comments pertaining to the Age-Appropriate Design Code and age verification/digital identity systems above.

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

3.27. We would draw Ofcom's attention back to Media Smart, the media literacy programme initially focused on advertising. Its objective remains to provide children and young people with the tools to help them understand and interpret content they see so that they are able to make informed choices. Advertisers pay for the materials created, agencies support broadcast creative, media give airtime, and the industry platforms Media Smart at events. Media Smart content has included material on body image, social media, influencers, and online advertising. We will continue to support these efforts to promote media literacy among the next generation and online harms regimes are put in place.

The logo consists of a dark blue speech bubble with a tail pointing towards the bottom left. Inside the bubble, the words "One Voice" are written in a bold, dark blue, sans-serif font, with "One" on the top line and "Voice" on the bottom line.

**One
Voice**