

# Written evidence submitted by ISBA

## Response to DCMS Sub-Committee on Online Harms and Disinformation

### *Online safety and online harms*

September 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. Consultation Response**

- 2.1. ISBA has welcomed and encouraged the development of debate and legislation around online harms over the past several years. We have supported government in its desire to deliver a “world-leading package of online safety measures” – an ambition set out in the online harms White Paper.
- 2.2. In this vein, we were pleased to see the Government’s response to the consultation on the White Paper, and the publication of the Draft Online Safety Bill. We remain of the view that while there are huge opportunities inherent in the development of the digital economy, we also face serious challenges to individual and collective safety. Meeting them is a global task and one in which the advertising and marketing industry must, will, and is playing its full part.
- 2.3. We have previously advocated for proportionate regulation of the major digital platforms, based on the principles of an effectively, fairly-funded, and collective regulatory environment; a transparent and independent evidence base of clear, understandable information for advertisers and consumers; and redress through an independent arbitration process backed by co-regulation. This approach was drawn from our perspective as the trade body for brand advertisers in the UK, and the need for responsible advertisers to have responsible digital partners.
- 2.4. In short, marketers need to have confidence in the content policies of platforms, and be assured that they offer consistent levels of protection (and that these are being adhered to) before they can decide whether to invest in those channels. The work which is going on with this legislation – and within industry, via international efforts of which ISBA is a key part (on which more below) – is key to developing this confidence.

2.5. Although we had argued for a new and dedicated regulator, we support the designation of Ofcom to these duties, albeit that we hope it will be equipped with the necessary funds and expertise to support it to carry out the role effectively. In this response, and as the debates on the Bill proceed, we hope that we will see Ofcom provided with the framework for overseeing an effective, fairly-funded, and collaborative effort to prevent online harms – with commonly held principles and codes of conduct supporting systemic transparency and accountability.

*How has the shifting focus between ‘online harms’ and ‘online safety’ influenced the development of the new regime and draft Bill?*

2.6. We welcome the increased role and responsibility which Ofcom will receive through Part 4 of the Bill to improve the media literacy of the public as the end-users of online services – including the need to counter misinformation and disinformation, and to commission or encourage initiatives which improve media literacy rates; as well as encouraging regulated service providers to develop tools which can improve such literacy, and develop products which can help the public identify the types of material they are seeing and interrogate it.

2.7. Our industry continues to play its part in developing media literacy, especially among young people. Media Smart is the advertising industry’s education programme, the mission of which is to ensure that young people in the UK can confidently navigate the media they consume – including being able to identify, interpret, and critically evaluate all forms of advertising.

2.8. Media Smart creates free media and digital literacy resources for teachers, parents, and youth organisations working with 7-16 year olds. Past education resources have focused on social media, digital advertising, influencer marketing, data, and piracy.<sup>1</sup>

2.9. We will continue to support this week as an industry, and work with Ofcom as it undertakes its expanded responsibilities. We would note that the shift in emphasis in the Bill away from online ‘harms’ to online ‘safety’ should not result in a lack of recognition from platforms of their continued responsibility to actively prevent harm on their networks by creating the right structures, checks, and balances. Nor should it mean that individuals absolve themselves of their own responsibilities both not to commit harm themselves, or indeed to become more media literate.

2.10. We hope that Ofcom will be supported with the resources it needs in order to be able to fulfil this role, and promote media literacy to users of online services of all ages and backgrounds.

*Is it necessary to have an explicit definition and process for determining harm to children and adults in the Online Safety Bill, and what should it be?*

2.11. The draft Bill would seem to give broad definitions of what might constitute harm, while leaving scope for Ofcom as the regulator to interpret the legislation and issue further guidance, ruling on individual cases and thereby setting precedent. This is ground-breaking legislation and among the first of its kind in the world. It will almost certainly therefore require fresh iterations and updates. With that in mind, it would seem prudent to design the legislation in as futureproof a way as possible, allowing

---

<sup>1</sup> Resources and information are available at <https://mediasmart.uk.com/>.

for adaptability on the part of the regulator and government, and an ability to respond to a rapidly-changing online environment.

- 2.12. In considering the specifics of what constitutes an online harm, the advertising industry has been convening internationally in an effort that unites marketers, media agencies, media platforms, and industry associations. The Global Alliance for Responsible Media (GARM) was established by the World Federation of Advertisers in 2019, and aims to safeguard the potential of digital media by reducing the availability and monetisation of harmful content online. ISBA is a member of the GARM Steering Committee, and we see this work as essential to creating a safer digital media environment that enriches society through content, communications, and commerce.
- 2.13. One of the first steps in safeguarding the positive potential for digital is to provide platforms, agencies, and marketers with the framework with which to define safe and harmful content online. One cannot address the challenge of harmful online content if one is unable to describe it using consistent and understandable language.
- 2.14. GARM has developed and will adopt common definitions to ensure that the advertising industry – from brands and trade bodies to large platforms such as Facebook and Google – is categorising harmful content in the same way across the board. Eleven key categories have been identified in consultation with experts from GARM's NGO Consultative Group. Establishing these standards is the essential foundation needed to stop harmful content from being monetised through advertising. Individual GARM members will adopt these shared principles in their operations, whether they are a marketer, agency, or media platform; and platforms including Facebook, YouTube and Twitter are among those who have committed to the framework for defining harmful content that is inappropriate for advertising. They have also agreed to collaborate with a view to monitoring industry efforts to improve in this area.
- 2.15. 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. The Suitability Framework (Fig. 2) lists sensitive content which may be appropriate for advertising, when that advertising is supported by proper controls.
- 2.16. 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 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.

*Does the draft Bill focus enough on the ways tech companies could be encouraged to consider safety and/or the risk of harm in platform design and the systems and processes that they put in place?*

- 2.17. The GARM-led international effort aims to standardise definitions and classifications of harmful content so that it can be more consistently identified by machines and humans. In this way, efforts to improve brand safety or suitability in a programmatic environment can be made more effective and predictable.

<b>CONTENT CATEGORY</b>	<b>BRAND SAFETY FLOOR – Content not appropriate for any advertising support</b>
Adult & Explicit Sexual Content	<ul style="list-style-type: none"> <li>• Illegal sale, distribution, and consumption of child pornography</li> <li>• Explicit or gratuitous depiction of sexual acts, and/or display of genitals, real or animated</li> </ul>
Arms & Ammunition	<ul style="list-style-type: none"> <li>• Promotion and advocacy of Sales of illegal arms, rifles, and handguns</li> <li>• Instructive content on how to obtain, make, distribute, or use illegal arms</li> <li>• Glamorization of illegal arms for the purpose of harm to others</li> <li>• Use of illegal arms in unregulated environments</li> </ul>
Crime & Harmful acts to individuals and Society, Human Right Violations	<ul style="list-style-type: none"> <li>• 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.),</li> <li>• Harassment or bullying of individuals and groups</li> </ul>
Death, Injury or Military Conflict	<ul style="list-style-type: none"> <li>• Promotion, incitement or advocacy of violence, death or injury</li> <li>• Murder or Willful bodily harm to others</li> <li>• Graphic depictions of willful harm to others</li> <li>• Incendiary content provoking, enticing, or evoking military aggression</li> <li>• Live action footage/photos of military actions &amp; genocide or other war crimes</li> </ul>
Online piracy	<ul style="list-style-type: none"> <li>• Pirating, Copyright infringement, &amp; Counterfeiting</li> </ul>
Hate speech & acts of aggression	<ul style="list-style-type: none"> <li>• 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.</li> </ul>
Obscenity and Profanity, including language, gestures, and explicitly gory, graphic or repulsive content intended to shock and disgust	<ul style="list-style-type: none"> <li>• Excessive use of profane language or gestures and other repulsive actions that shock, offend, or insult.</li> </ul>
Illegal Drugs/Tobacco/e-cigarettes/Vaping/Alcohol	<ul style="list-style-type: none"> <li>• 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</li> <li>• Promotion and advocacy of Tobacco and e-cigarette (Vaping) &amp; Alcohol use to minors</li> </ul>
Spam or Harmful Content	<ul style="list-style-type: none"> <li>• Malware/Phishing</li> </ul>
Terrorism	<ul style="list-style-type: none"> <li>• Promotion and advocacy of graphic terrorist activity involving defamation, physical and/or emotional harm of individuals, communities, and society</li> </ul>
Debated Sensitive Social Issue	<ul style="list-style-type: none"> <li>• Insensitive, irresponsible and harmful treatment of debated social issues and related acts that demean a particular group or incite greater conflict;</li> </ul>

**Fig. 1. GARM Brand Safety Floor**

- 2.18. Our belief is that, despite their efforts so far, there is scope for further investment by the large platforms and improvement in the technologies that are able to automatically identify and understand the nuanced meaning of text and images in real time. There should be consistency in the application of minimum acceptable thresholds for their adoption. The development of these technologies has lagged behind the development of core revenue-driving systems, as they dilute profitability and were not identified as critical requirements during the platforms' earlier growth phases.
- 2.19. The undoubted scale of the amount of content on platforms – particularly that which is user-generated – makes real-time identification challenging. 100% identification of problematic content may be an unrealistic goal, but some of the platforms' own transparency reporting suggests that significant improvements can be made by the application of machine learning and human moderation in combination. As we have seen, major platforms do now report prevalence (in the form of the number of views of harmful content as a percentage of total views) at a global level. It is an ongoing aim of GARM to enable reporting of problematic and unacceptable content at a national level, and for it to be possible to identify surge incidents (such as the hatred directed at England's footballers after the Euro 2020 final, or interference in particular elections) and the action taken by platforms to deal with them.

CONTENT CATEGORY	High Risk	Medium Risk	Low Risk
Adult & Explicit Sexual Content	<ul style="list-style-type: none"> <li>Suggestive sexual situations requiring adult supervision/approval or warnings</li> <li>Full or liberal Nudity</li> </ul>	<ul style="list-style-type: none"> <li>Dramatic depiction of sexual acts or Sexuality issues presented in the context of entertainment</li> <li>Artistic Nudity</li> </ul>	<ul style="list-style-type: none"> <li>Educational, Informative, Scientific treatment of sexual subjects or sexual relationships or sexuality</li> </ul>
Arms & Ammunition	<ul style="list-style-type: none"> <li>Glamorization /Gratuitous depiction of illegal sale or possession of Arms</li> <li>Depictions of sale/use/distribution of illegal arms for inappropriate uses/harmful acts</li> </ul>	<ul style="list-style-type: none"> <li>Dramatic depiction of weapons use presented in the context of entertainment</li> <li>Breaking News or Op-Ed coverage of arms and ammunition</li> </ul>	<ul style="list-style-type: none"> <li>Educational, Informative, Scientific treatment of Arms use, possession or illegal sale</li> <li>News feature stories on the subject</li> </ul>
Crime & Harmful acts to individuals and Society, Human Right Violations	<ul style="list-style-type: none"> <li>Depictions of criminal/harmful acts or violation of human rights</li> </ul>	<ul style="list-style-type: none"> <li>Dramatic depiction of criminal activity or human rights violations presented in the context of entertainment</li> <li>Breaking News or Op-Ed coverage of criminal activity or human rights violations</li> </ul>	<ul style="list-style-type: none"> <li>Educational, Informative, Scientific treatment of crime or criminal acts or human rights violations</li> <li>News feature stories on the subject</li> </ul>
Death, Injury or Military Conflict	<ul style="list-style-type: none"> <li>Depiction of death or Injury</li> <li>Insensitive and irresponsible treatment of military conflict, genocide, war crimes, or harm resulting in Death or Injury</li> <li>Depictions of military actions that glamorize harmful acts to others or society</li> </ul>	<ul style="list-style-type: none"> <li>Dramatic depiction of death, injury, or military conflict presented in the context of entertainment</li> <li>Breaking News or Op-Ed coverage of death, injury or military conflict</li> </ul>	<ul style="list-style-type: none"> <li>Educational, Informative, Scientific treatment of death or injury, or military conflict</li> <li>News feature stories on the subject</li> </ul>
Online piracy	<ul style="list-style-type: none"> <li>Glamorization /Gratuitous depiction of Online Piracy</li> </ul>	<ul style="list-style-type: none"> <li>Dramatic depiction of Online Piracy presented in the context of entertainment</li> <li>Breaking News or Op-Ed coverage of Online Piracy</li> </ul>	<ul style="list-style-type: none"> <li>Educational, Informative, Scientific treatment of Online Piracy</li> <li>News feature stories on the subject</li> </ul>
Hate speech & acts of aggression	<ul style="list-style-type: none"> <li>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</li> </ul>	<ul style="list-style-type: none"> <li>Dramatic depiction of hate speech/acts presented in the context of entertainment</li> <li>Breaking News or Op-Ed coverage of hate speech/acts</li> </ul>	<ul style="list-style-type: none"> <li>Educational, Informative, Scientific treatment of Hate Speech</li> <li>News features on the subject</li> </ul>
Obscenity and Profanity, including language, gestures, and explicitly gory, graphic or repulsive content intended to shock and disgust	<ul style="list-style-type: none"> <li>Glamorization /Gratuitous depiction of profanity and obscenity</li> </ul>	<ul style="list-style-type: none"> <li>Dramatic depiction of profanity and obscenities presented in the context of entertainment by genre</li> <li>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</li> </ul>	<ul style="list-style-type: none"> <li>Educational or Informative, treatment of Obscenity or Profanity</li> <li>News feature stories on the subject</li> </ul>
Illegal Drugs/Tobacco/e-cigarettes/Vaping/Alcohol	<ul style="list-style-type: none"> <li>Glamorization /Gratuitous depictions of illegal drugs/abuse of prescription drugs</li> <li>Insensitive and irresponsible content/treatment that encourages minors to use tobacco and vaping products &amp; Alcohol</li> </ul>	<ul style="list-style-type: none"> <li>Dramatic depiction of illegal drug use/prescription abuse, tobacco, vaping or alcohol use presented in the context of entertainment</li> <li>Breaking News or Op-Ed coverage of illegal drug use/prescription abuse, tobacco, vaping or alcohol use</li> </ul>	<ul style="list-style-type: none"> <li>Educational, Informative, Scientific treatment of illegal drug use/prescription abuse, tobacco, vaping or alcohol</li> <li>News feature stories on the subject</li> </ul>
Spam or Harmful Content	<ul style="list-style-type: none"> <li>Glamorization /Gratuitous depiction of Online Piracy</li> </ul>	<ul style="list-style-type: none"> <li>Dramatic depiction of Spam or Malware presented in the context of entertainment</li> <li>Breaking News or Op-Ed coverage of Spam or Malware</li> </ul>	<ul style="list-style-type: none"> <li>Educational, Informative, Scientific treatment of Spam or Malware</li> <li>News feature stories on the subject</li> </ul>

**Fig. 2. GARM Suitability Framework**

2.20. Learning from incidents such as the Euro 2020 final is crucial. There must be continual improvement and raising of minimum acceptable standards – standards which the new regulator must set. It is important that we are able to evaluate the true impact of harmful content – for example, how many people actually saw the racist abuse directed at our footballers? What were the lessons learned by the platforms in terms of how their systems identified and/or removed that abuse? Are there remedial future steps that they can take? And, crucially: how will the platforms report and account for this information publicly; and what independent verification will there be of it? By way of an example, Facebook enables a third-party audit of its Community

Standard Enforcement Report (CSER), which is carried out by one of the so-called 'Big Four' accounting firms. This is welcome, but in our view, it is clear that Ofcom should now set the standards for such audits – setting minimum standards, and raising the bar for action and compliance.

*What are the key omissions to the draft Bill, such as a general safety duty or powers to deal with urgent security threats, and how could they be practically included without compromising rights such as freedom of expression?*

- 2.22. We have noted the debate around the absence from the draft Bill on measures to tackle advertising fraud. This is a serious issue which puts consumers at risk of harm, undermines the credibility of industry, and poses a threat to the security of paid-for online advertising.
- 2.23. We are aware of the Government's concern about this subject, and equally of their belief that this issue should be tackled in the round along with other considerations about paid-for advertising as part of DCMS's Online Advertising Review. Consistent with our previous advocacy for holistic policymaking and for settling on the overall architecture of online advertising policy before dealing with individual policy issues, we believe that this is the right approach.
- 2.24. This is not to diminish the concerns raised by stakeholders, including for instance MoneySavingExpert, that scams being pushed through paid-for advertisements which appear in internet search results, promoted posts on social media, and online dating profiles represent a real danger. There are loopholes which can be exploited by online scammers and organised crime. ISBA fully supports efforts to tackle this ad fraud, in the context of ensuring that paid-for online advertising is trusted, transparent, and accountable to regulators and consumers.

*What are the lessons that the Government should learn when directly comparing the draft Bill to existing and proposed legislation around the world?*

- 2.25. As Ministers have stated, the Online Safety Bill is among the first pieces of legislation globally to deal with internet regulation and online harms in a systemic way. Government has proceeded slowly – to some criticism – although we recognise that this can allow for collaboration with industries such as our own. We hope that government has been working with partners internationally, just as our industry has, on creating shared sets of standards and definitions – bearing in mind that we are dealing here with platforms which are multinational, and global in their reach.
- 2.26. As with GDPR before it, we believe that the major tech firms will set their global standards based on leadership positions taken by regulators in significant markets. Ofcom and the UK clearly fit this description. The Online Safety Bill has the potential to have such a global impact, and – although government can obviously only legislate for UK jurisdiction – the potential for wider, if not global, implementation should be borne in mind.