Your response

Volume 2: The causes and impacts of online harm

Ofcom’s Register of Risks

Question 1:

i) Do you have any comments on Ofcom’s assessment of the causes and impacts of online harms?

Response: There appear to be a few ‘chinks’ in the regulatory ‘armour’ proposed by OSA2023 and Illegal Harms consultation materials. These chinks are listed below. Interested in hearing how Ofcom would unpack all this in the interests of effective enforcement. Effective enforcement would embrace both the detection of offences where they are being committed and also the assessment that a Service has a clean bill of health where no offences are believed to be occurring. Would Ofcom consider using the Codes of Practice to close up any regulatory chinks? This would provide clarity for both regulator and regulated.

1). Section 1(1). The active bit for an enforcement inspector is in bold: “*This Act provides for a new regulatory framework which has the general purpose of* ***making the use of internet services*** *regulated by this Act* ***safer for individuals*** *in the United Kingdom ….*”. Some questions arise:

a. What are the criteria for ‘safer’?

b. How will the regulated know they have done enough to satisfy the regulator?

c. Where is the definition of ‘safer’? The Act contains no definition. There are tentative proxies (shown in consultation materials which cannot be put into the Act) that Ofcom hope will deliver on ‘safer’.

i. Ref “Ofcom’s approach to implementing the Online Safety Act” Page 7; Under ‘We expect change’ refers to: “we **anticipate** the Act will ensure people in the UK are safer online by delivering four outcomes”. What is the basis of the anticipation in the absence of a definition of ‘safer’?

ii. Ref “Ofcom’s approach to implementing the Online Safety Act” Page 7; Under Key Outcome 4: refers to: “*ability of regulation to deliver a safer life online*”. What is the basis of this claim in the absence of a definition of ‘safer’?

iii. Vol 2, page 75, para 6C.169, refers to: “***Young people*** *using these features may* ***believe*** *that their images* ***are safer*** *by sharing them in this format, in that there will be no permanent record of them, however the* ***evidence suggests*** *that users can deploy* ***tactics to circumnavigate this feature****.*” There is some indication of what ‘safer’ looks like, but nothing that can be tested. More worryingly Ofcom is indicating that image sharing may become less safe because controls can be by-passed or defeated. This would be an important area for expansion in a Code of Practice (COP).

iv. Vol 3, page 19, para 8.72, refers to: “***we think*** *it likely that where* ***such costs*** *are incurred, they are* ***likely to be proportionate*** *as there are* ***likely benefits*** *from* ***users being safer from illegal content.***” SO it seems the possibility of ‘safer’ being achieved depends on the Accountable Person being appointed, but Ofcom only ‘thinks’ that safer might be achieved. This diffidence appears to be at odds with the strong assertion in Section 1(1). AND it seems highly likely that cost considerations would put Service providers off and there is no mechanism that can compel the provider to address ‘safer’ with an appointed person, BECAUSE the usual means of cost benefit analysis in relation to risks does not appear to be intended to be deployed.

v. Vol 4 has 12 occurrences of ‘safer’. Two (2) re-iterate purpose of the Act. Four(4) are to do with ‘feeling’ safer. Which suggests a Subjective approach rather than an Objective approach is being adopted by Ofcom. Clearly online safety is a challenge. A regulator does need to be able to define some red lines so that everyone knows when enforcement is likely or not. The more subjective the regime, the less likely bad outcomes will be challenged early enough.

vi. Vol 5, page 49, para 26.260, refers to: “*We are committed to reducing harm from cyberflashing as part of our wider effort to make the **online space safer** for women and girls*.”. This is a fantastic and desirable aspiration [for the online space] and very much akin to the Health and Safety at Work etc. Act 1974 (HSWA74) focus on workplace safety. It is a much bigger remit than “***making the use** of **internet services** regulated by this Act **safer for individuals***.”. As a point of detail, the online space needs to be safer for ALL people regardless of gender or sexual orientation.

vii. Vol 6, page 25, “What this chapter is about” refers to: “*This chapter sets out our approach to **supervision of** a small subset of the highest reach or **highest risk services** in scope of the Online Safety Act. Supervision **will help ensure** that **these services** have appropriate systems and processes to achieve the key outcomes intended by the Act to **make life safer online for people across the UK***.” This too is aspirational like Vol 5, focusing on ‘life online’ – which we can take to mean ‘online space’ – and not just the component parts [Internet Services] that make up that overall system of an online space. The Act however in Section 1(1) is only looking for the components to be safer, whilst internal to Ofcom there seems to be a desire to embrace the whole domain of the online space; exactly as the public and politicians would expect the online safety regulator to do.

d. The above suggests that a Service could be deemed ‘safer’ by the Provider if it was less bad (and/or less risky?) after some measures applied. As this is the only occurrence of the word ‘safer’ in the Act, not possible to see how any of the other clauses are connected to this statement of purpose or objective.

e. The Act (on its own without the regulator chinks closed in the Codes of Practice) is inadequate to police online safety, as there is no requirement for ensuring that Service(s) are fit for purpose; which can then be assessed as safe or not.

2). Section 1(2). The active bit for an enforcement inspector is in bold: “To achieve that purpose, **this Act** (among other things)— (a) **imposes duties** which, in broad terms, **require providers of services** regulated by this Act **to identify, mitigate and manage the risks of harm** (including risks which particularly affect individuals with a certain characteristic) **from**— (i) **illegal content and activity**, and (ii) **content and activity that is harmful to children**, and …”.

a. At face value it might be possible to see that 1(2) says it achieves the purpose at 1(1) by imposing duties. In order for the duties of identifying, mitigating and managing the risks of harm to satisfy 1(1) the enforcement inspector needs to be able to satisfy themselves that the actions of the Service provider have ‘made the use of internet services … safer for individuals’. As there is no assessment process for what is ‘safer’, then it appears Section 1(2) does not provide any leverage for Ofcom to urge the service provider to any remedial action.

b. Section 1(2) narrows the focus of the ‘risk of harm’ lens, to that which involves ‘illegal content and activity, and (ii) content and activity that is harmful to children’. Two observations arise.

i. ONE that legal content and activity could be excluded from Ofcom’s mandate even if it gives rise to harm; which surely is not what the public want and need.

ii. TWO with so many ‘and’s in the clause – rather than ‘or’s – it could mean that any one of the four situations not being satisfied would rule out Ofcom’s mandate again; simply because of how the law would be evaluated in court, especially by the Defence Team. The four situations being (illegal content; illegal activity; content harmful to children; activity harmful to children.

3). Section 1(3)(a) [services regulated by this Act] are—safe by design.

a. There is only one occurrence of ‘safe by design’ in the whole Act. As there is no assessment process for what is ‘safe by design’ then it appears Section 1(3)(a) is an end in itself which appears not to be linked to the stated Purpose of the Act in Section 1(1).

4). Section 1(3)(b)(i) [services regulated by this Act are designed and operated in such a way that] a higher standard of protection is provided for children than for adults.

a. As there is no assessment process for what ‘standard of protection’ is provided for either children or adults, then there can be no comparison. This means that services are being asked to evaluate a situation they are not legally required to gather information on.

5). Section 1(3)(b)(ii) [services regulated by this Act are designed and operated in such a way that] users’ rights to freedom of expression and privacy are protected,

a. Typically regulators are asked to focus on harms and potential for things to go wrong. This requires the regulator to assess the ‘goods’ or delivery of benefits of freedom of expression and privacy AND to ensure that these are protected whilst assessing potential for harm. This is a complex interaction of two totally different requirements that manifest in different ways and whose mitigations operate differently. It is not clear how the Service provider or regulator are expected to process this in a cost effective and meaningful manner.

6). Section 1(3)(b)(iii) [services regulated by this Act are designed and operated in such a way that] transparency and accountability are provided in relation to those services.

From Section 178(3)(a)(iii) we learn that ‘transparency and accountability’ is to users by service providers. Also that this quality of ‘transparency and accountability’ is required to be reported to the Secretary of State; yet it appears there is no process of assessment. So it is not clear what is required of the service provider to demonstrate this, nor what the User would look for.

ii) Do you think we have missed anything important in our analysis? Please provide evidence to support your answer.

Response: Yes: In short; Ofcom’s Risk Management proposals however good they may be in their own right, are not being required to serve the most important objective (MIO) in Section 1(1) of OSA2023. Without a definition of ‘safer’ the Risk Management activities proposed will end up serving themselves and evolve into being a Tick Box regime, however much Ofcom, politicians and society would wish them not to become [tick box exercises]. To complicate matters there are five (5) different interpretations of ‘safer’ deducible within the Illegal Harms consultation materials; making the need for definition(s) of ‘safer’ more urgent. Is Ofcom proposing to use the Codes of Practice to set out the definitions, criteria, assessments and processes in relation to Sections 1(1), 1(2) and 1(3)? If so this would transform OSA2023 into a powerful tool able to deliver fit for purpose and safe online experiences for all users.

1. Section 1(1) of OSA2023 declares the most important objective of [synopsis]: "*Internet Services are safer for individuals to use*". This is a Product use requirement, very similar to HSWA74 Section 6. However this is as far as the Act gets, so all the Risk Management activity automatically defaults to failures and responses to those failures. Written as it is, this clause means it is not possible to be pre-emptive or Anticipatory in action, which is necessary to control harm and ensure the Internet is fit for purpose and safe. In other words, harms are not prevented – they take place, people are harmed. Meanwhile the provider can fail to achieve 'safe' or 'safer' in respect of a particular aspect of service 'use'. Reference to the wider context of the 'online space' within which that use is taking place, would provide a criteria from which to assess if remedial measures are needed.
2. There are five interpretations in OSA2023 of 'safer' deduced from Consultation materials, none of which can be, or are required to be tested by regulator or the regulated. These are:
    a. Type 1 Safer; Placeholder or general label, with no definition assigned.

    i). all the references to the "Purpose of the Act is to make the use of regulated internet services safer for individuals in the United Kingdom."

    ii) Vol 5, Page 49, para 26.260; 'safer' = related to a whole online space, of which cyberflashing is part. Is there a list of all the different parts of 'safer' that would make up this whole so everyone can see what OSA2023 Section 1(1) is mandated to address?

    b. Type 2 Safer; Acute Safety Control by absolute elimination of exposure to Harm or Hazard.

    i). Vol 2, Page 75, para 6C.169; 'safer' = absolute elimination of exposure to deleted messages.

    ii). Vol 3, Page 19, para 8.72; 'safer' = absolute elimination of exposure to illegal content. (although in this instance it is not clear precisely what 'safer from illegal content' means, and how the absence or presence of safety would be assessed.)

    iii). Vol 4, Page 238, para 18.27; first word 'safer' = implication of elimination of exposure to harm in general. (although in this instance it is not clear precisely what 'safer settings [for child users]' means, and how the absence or presence of safety would be assessed.)

    c. Type 3 Safer; Psychological Perception of being Safer or not Safer

    i) Vol 4, Page 238, para 18.27; second word 'safer' = a subjective assessment of safety by a 3rd party or the user themselves, without any evidential route to evaluating the truth or otherwise of this conclusion. (in this instance it is not clear how Ofcom has reached this conclusion that: [default settings for child users] 'would tend to make children overall safer in their online experiences'.)

    ii) Vol 4, Page 256, para 18.105; 'safer' = a subjective assessment of safety of decisions being made by users. There is also a companion element of Type 2 safer needing to be exercised by the designer of the service, to encourage user choices that lead to safe outcomes.

iii) Vol 4, Page 257, para 18.107(a); 'safer' = a subjective assessment of safety in relation to interactions with other people/users. This is part of a complex interaction of people and information prompts, which is not clear how this would make a situation safer.

iv) Vol 4, Page 259, para 18.120; 'safer' = a subjective assessment of safety after a precursor action. This scenario has Ofcom claiming "*child users feeling* ***safer offline*** *after having taken action on the platform*"; not clear how this claim can be validated. It also appears that the Offline space is outside Ofcom's jurisdiction of the Online space.

v) Vol 4, Page 259, para 18.121; 'safer' = a subjective assessment of safety after a precursor action. This appears to be an unnecessarily ambiguous and low performance outcome in respect of safety. If the child user is being provided with 'clear information about offender threatening action' then they, and their guardians, surely would want to know they are actually safer rather than just 'feel safer'. The child user ought to be benefiting from Type 2 Safer as a precursor to Type 3 safer. And in any event the framing of this scenario is Lagging in respect of safety, because the use of the word offender implies that an offence has already taken place, so harm has not been prevented, and neither has exposure to that harm been prevented.

d. Type 4 Safer; Emotional Anticipation & Expectation of being Safer in the future

i) Vol 4, Page 259, para 18.122; 'safer' = an expectation of being able to feel safer, as a consequence of Type 2 safer interventions to block exposure to harm. This is complex to implement and regulate as the expression of Type 4 safer is often easier than the embedding of Type 2 Safer in platform systems. Feeling safer may be dangerous in the online world if the conditions are not fit for purpose and able to deliver an experience that is safe. It is one thing to 'feel' safer, it is another to really be safer, as testified by several offline child abuse cases. System Safety design needs to address the potential for the False Positives of feeling safer when in fact the user should feel less safe.

ii) Vol 4, Page 262, para 18.138(c); 'safer' = a presumed expectation of feeling safer online having delivered information to child users. The promise of a Type 2 level safer, that can yield genuine Type 3 safer in the child, can easily be undermined by malicious actors and/or ineffective absorption of information. This apparent strength can hide a weakness in safety provision, in plain sight.

e. Type 5 Safer; Deemed Economically Safer System

i) Vol 4, Page 291, para 20.40; 'safer' = is a Type 3 Safer moderated by the economics of the platform as a consequence of blocking reducing user engagement. Paras 20.1 to 20.42 present the potentials for exposure to harm. These paras also present the primary mechanism of blocking users, to provide shielding against harms. Para 20.40 appears to be focusing on the difficulty and impacts of individual vs mass blocking, and expressing this as a cost to platforms in implementation challenges & effort and/or changes in revenue. This cost-led argument by default puts safety as a secondary consideration, which appears to go against Section 1(1) of OSA2023. The platform systems need to be more nuanced and intelligent to present to users the best proportion of individual and mass blocking options, in a dynamic manner. Yes this will require advanced programming skills, but this is feasible and necessary in the primary interests of securing fit for purpose and safe (FFP&S) experiences for all users. Securing FFP&S must be the primary objective for the regulator rather than the lower level objective of users 'feeling safer online'.

3. The Consultation material expands Ofcom interests into the whole online space (Vol 5 and 6) – which is applauded as being more in line with HSWA74 – whilst the Act stays focused on the component parts [internet services] that make up that online space. What exactly is in Ofcom's mind when it is discharging its duties? What is Ofcom expecting its regulatory staff to have in mind; The big picture or the parts of the big picture?
4. It is not evident how OSA2023 Section 1(3) supports 1(2) and in turn not evident how 1(2) supports 1(1). Is Ofcom planning to provide guidance and instruction in its Codes of Practice?
5. The opening paragraph of the Act says it is about 'Communications Offences'. The focus on 'offences' narrows the regulatory lens to failures and what has gone or could go wrong. Offences are AGAINST the common good. The Act does not appear to set itself out to be in pursuit of outcomes that are FOR the common good. This seems to be a very narrow focus compared to the online space or the parts being 'safer'. Commentary in the COPs would be of great help to appreciate the scope of regulation Ofcom has in mind.
6. It is unclear what a new entrant regulator for Ofcom, based on OSA2023 and Consultation materials, would know for sure what they were allowed to do in the field by way of Effective Enforcement. Does Ofcom have materials planned for regulatory staff to ensure they know what is expected of them, and how they can communicate this to the regulated and public at large?

iii) Is this response confidential? (if yes, please specify which part(s) are confidential)

Response: No

Question 2:

i) Do you have any views about our interpretation of the links between risk factors and different kinds of illegal harm? Please provide evidence to support your answer.

Response: This is related to Q1(i) and Q1(ii) responses above. Any linkage between Risk Factors and Illegal harms need to be established in the context of Section 1(1) with the higher goal of 'use of services being safer for individuals'.

ii) Is this response confidential? (if yes, please specify which part(s) are confidential)

Response: No