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 are a number of references to children "abusing themselves" and "abusing others". For example, "performing abusive acts on others" within the definition of contact child sexual abuse, and in the summary of livestreaming. Children do not abuse themselves, they are abused. Where children are groomed, manipulated or coerced into performing harmful sexual acts on other children they are still being abused and we need to ensure the language used does not portray them as the perpetrator.

We are concerned that the impact of 'online only' child sexual abuse is minimised throughout, with the emphasis on harm only being caused if it leads to contact abuse. We know that online abuse is just as harmful as contact child sexual abuse, and in some cases more so - "TACSA is no less impactful than offline-only sexual abuse." (NSPCC Authors: Dr Catherine Hamilton-Giachritsis, Dr Elly Hanson, Dr Helen Whittle and Professor Anthony Beech Published: 2017, Impact of online and offline child sexual abuse: "Everyone deserves to be happy and safe" | NSPCC Learning.) We support many children and young people whose experiences reflects this. We have provided below an anonymised case study written by a mother which highlights the impact of 'online only' abuse. We feel that this should be given greater emphasis within the proposals to reflect that this does lead to a severe and often lifelong impact for the victim.

••

[CONFIDENTIAL case study ✂]

Response:

P. 8 Summary: "A particularly clear example of this is online grooming, which can result in contact sexual abuse and cause lifelong negative psychological impacts including, loss of confidence, aggression, feelings of self-blame and lack of personal trust, as well as an increased risk of self-harm."

We feel aggression should not be included in this list of impacts of online grooming. "Aggression is how we choose to behave in response to a stimulus and involves forcibly and harmfully attempting to get others to behave differently. Exerting ourselves over others in any way is an unacceptable way of dealing with our emotions and can be bullying and harmful. When anger feelings are confused with aggressive behaviours it can result in any healthy anger being denied." (https://www.southdownspsychotherapy.co.uk/anger-vs-aggression-whats-the-difference/#:~:text=Anger%20is%20an%20acceptable%20feeling,confusion%20between%20the%20two%20emotions.) Having aggression as an impact of online grooming, phrased in this way, negates the trauma response that victims exhibit or feel. There is no choice in a trauma response.

Summary P.29 – "In this chapter we explore features that U2U services can use to help users manage the risk of being exposed to illegal content."

Why is it on the user to manage the risk? It is the service's responsibility to ensure that there is no illegal content on their platform, and therefore it is for them to ensure users are not exposed to this type of content. We would suggest changing the language here to reflect the responsibility being on the platform rather than the user, although we recognise that giving the user autonomy as a second line of defence is a good thing.

Vol.2 P.3 – "The functionalities we describe above are not inherently bad and have important benefits. End-to-end encryption plays an important role in safeguarding privacy online."
We feel that there is an important point around privacy that has been missed, namely that E2EE does not always protect privacy, and in fact for victims and survivors of image-based CSA it denies their right to privacy. E2EE does not enable services to detect and remove child sexual abuse from the encrypted parts of their service. Enabling known child sexual abuse material to circulate freely means that the children in those images are being continually revictimized with no control and no mechanism for their images to be detected and removed.

A further point around E2EE is that, while we recognise the limitations of the Online Safety Act, we strongly feel that service providers should be compelled to ensure that children are protected in all areas of their platforms, including private messages and E2EE areas of their services. We do not feel that these proposals sufficiently address the dangers posed. At numerous points there is reference to how perpetrators operate and seek E2EE services to groom and abuse children, and yet there is a lack of proposed measures to address services who are knowingly implementing E2EE without appropriate safeguards to ensure children are protected or to remove. Platforms will not be truly safe for children and young people, without also including the private and/or E2EE areas of platforms. Action to tackle grooming and CSAM will not be sufficient with half measures.

Vol.2 6c.113 - In these cases, the negative impacts on the child in the imagery are significant and extend to mental health problems and negative social repercussions, especially for girls, who will often experience bullying, harassment, social exclusion and victim-blaming."

We feel the harm caused through non-consensual image sharing is not truly reflected in the consultation. The impact of this type of abuse can be life-long and severe and sadly in some cases can result in children taking their lives. Whilst some research findings may indicate that girls suffer more frequent harm as a result of this type of abuse, we know boys are often do not report this type of abuse which can result in gender bias in research. We know through our direct work with children and families, that this type of abuse can and does have an equal amount of harm on boys as it does on girls though the numbers reported may be less.

Vol.2 6C.119 – "Being sexually abused has a profound and long-lasting impact on a child. The existence, sharing and viewing of images and videos of that abuse can become a continuing source of trauma for victims and survivors of online CSAM. Many describe feeling constantly in fear, and vulnerable, because their abuse exists as a permanent record online which others can view."
We would add that trauma is increased by the complete lack of control over the images, which could reappear or be shared in any way at any time. Therefore, the abuse has no end date.

Vol.2 6C.120 – "Survivors can be re-victimised through reliving the experience of their sexual abuse if they see the material online. In a survey conducted by Canadian Centre for Child Protection (C3P), 69% of victims and survivors indicated that they worried constantly about being recognised, and almost a third (30%) had been identified online or in person by someone who had seen images of their abuse. Some victims and survivors reported being targeted and re-victimised by someone who had recognised them, including being propositioned or threatened. Victims and survivors also describe suffering from a heightened sensitivity to photos and cameras."
It is horrendous and completely unacceptable that some victims do see or are shown their abuse images. We would like to add that this is not the only way in which re-victimisation happens. A victim is re-victimised every time that material is viewed by anybody, regardless of whether the victim sees it or not. It is just not seeing that material online that causes harm, it is also knowing that the material may be online and may be seen by others for their own satisfaction. In addition, anxiety around images has increased since the introduction of AI generated CSAM which now provides perpetrators with the option to produce new synthetic images based on real abuse material. "In other words, real children who have been abused find themselves victimised again as offenders create new sexual imagery of them and distribute this online." (Lucy Faithfull Foundation https://www.lucyfaithfull.org.uk/featured-news/a-call-to-end-ai-generated-child-sexual-abuse.htm)

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

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

Response:

Volume 3: How should services assess the risk of online harms?

Governance and accountability

Question 3:
i) Do you agree with our proposals in relation to governance and accountability measures in the illegal content Codes of Practice?
Response:
ii) Do you think we have missed anything important in our analysis? Please provide evidence to support your answer.
• Response: We believe that organisations, in addition to being trained on approach must also receive training on understanding why they need to be compliant to counter bias.
iii) Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: No

Question 4:
i) Do you agree with the types of services that we propose the governance and accountability measures should apply to?
Response:
ii) Please explain your answer.
Response:
iii) Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Question 5:
i) Are you aware of any additional evidence of the efficacy, costs and risks associated with a potential future measure to requiring services to have measures to mitigate and manage illegal content risks audited by an independent third-party?
Response:
ii) Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Question 6:	
i)	Are you aware of any additional evidence of the efficacy, costs and risks associated with a potential future measure to tie remuneration for senior managers to positive online safety outcomes?
Response:	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:	

Service's risk assessment

Question 7:	
i)	Do you agree with our proposals?
Response:	
ii)	Please provide the underlying arguments and evidence that support your views.
Response:	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:	

Specifically, we would also appreciate evidence from regulated services on the following:

Question 8:	
i)	Do you think the four-step risk assessment process and the Risk Profiles are useful models to help services navigate and comply with their wider obligations under the Act?
Response:	
ii)	Please provide the underlying arguments and evidence that support your views.
Response:	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:	

Question 9:	
i)	Are the Risk Profiles sufficiently clear?
Response:	
ii)	Please provide the underlying arguments and evidence that support your views.
Response:	
iii)	Do you think the information provided on risk factors will help you understand the risks on your service?
Response:	
iv)	Please provide the underlying arguments and evidence that support your views.
Response:	
v)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:	

Record keeping and review guidance

Question 10:	
i)	Do you have any comments on our draft record keeping and review guidance?
Response:	
ii)	Please provide the underlying arguments and evidence that support your views.
Response:	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:	

Question 11:	
i)	Do you agree with our proposal not to exercise our power to exempt specified descriptions of services from the record keeping and review duty for the moment?
Response:	
ii)	Please provide the underlying arguments and evidence that support your views.
Response:	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:	

Volume 4: What should services do to mitigate the risk of online harms

Our approach to the Illegal content Codes of Practice

Question 12:
i) Do you have any comments on our overarching approach to developing our illegal content Codes of Practice?
Response:
ii) Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Question 13:
i) Do you agree that in general we should apply the most onerous measures in our Codes only to services which are large and/or medium or high risk?
Response: We are satisfied with the proposal for the most onerous measures to be applied to large and/or medium or high-risk services. We were particularly pleased to see recognition that small services will need to have more onerous measures applied where they are high risk for grooming/CSAM due to the harm that can be caused and to ensure users are protected from those risks even where they are using small services.
ii) Please provide the underlying arguments and evidence that support your views.
Response:
iii) Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Question 14:
i) Do you agree with our definition of large services?
Response: Yes
ii) Please provide the underlying arguments and evidence that support your views.
Response:
iii) Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Question 15:

i)   Do you agree with our definition of multi-risk services?
Response: Yes
ii)   Please provide the underlying arguments and evidence that support your views.
Response:
iii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Question 16:

i)   Do you have any comments on the draft Codes of Practice themselves?
Response:
ii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Question 17:

i)   Do you have any comments on the costs assumptions set out in Annex 14, which we used for calculating the costs of various measures?
Response:
ii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Content moderation (User to User)

Question 18:

i)   Do you agree with our proposals?
Response:
In relation to the prioritisation of content moderation, we strongly recommend that there be a time target for the taking down of CSAM. The longer that this content is available to be accessed on a service, even where it does not go viral, the more harm is caused to victims. Each time an image is viewed the child is revictimized. This has a significant impact on a victim's recovery journey, and they have the right to be assured that their abuse images will be removed as a priority. We would suggest that any CSAM reported should be removed within 24 hours at the very latest.
As a future consideration, we strongly feel that platforms should be pre-screening at the point of upload to ensure that CSAM does not appear on the platform in the first place. This will shift the burden of content moderation to be proactive rather than reactive and will result in less harm

caused, particularly to the victims and survivors in those images and videos. We see this as a gold standard safety by design measure that all services should begin working towards immediately and we would like to see this reflected in future proposals.

In relation to the training of content moderators we note that the proposal is for volunteer moderators to not receive training and materials. In relation to grooming and CSAM this is concerning for us as we would expect all content moderators, whether employed or voluntary, to have a sound knowledge and understanding of CSA to be able to recognise it when they see it, and to be able to take appropriate and timely action to remove content once discovered. We believe this proposal increases the risk of unconscious bias not being picked up and therefore this could lead to illegal content not being taken down. We would like to see the proposals include all content moderators.

ii) Please provide the underlying arguments and evidence that support your views.
Response:
iii) Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Content moderation (Search)

Question 19:
i)   Do you agree with our proposals?
Response: Yes
ii)   Please provide the underlying arguments and evidence that support your views.
Response:
iii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Automated content moderation (User to User)

Question 20:
i)   Do you agree with our proposals?
Response: Whilst we understand the reasons behind the user number thresholds, we urge that this be constantly reviewed because we believe that all services which are at high risk or medium risk for image-based CSAM, regardless of size, need to implement automated content moderation otherwise we risk creating spaces for people to disseminate CSAM freely without detection.
ii)   Please provide the underlying arguments and evidence that support your views.
Response: It is commonly known that perpetrators will seek out the most private places where they are least likely to be detected. We are already seeing this with E2EE (https://www.suojellaanlapsia.fi/en/post/tech-platforms-child-sexual-abuse). We believe that if there are smaller services available that do not use automated content moderation technology, they will become high risk, desirable places for perpetrators to operate.
iii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Question 21:
i)   Do you have any comments on the draft guidance set out in Annex 9 regarding whether content is communicated 'publicly' or 'privately'?
Response:
ii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Do you have any relevant evidence on:

Question 22:
i)   Accuracy of perceptual hash matching and the costs of applying CSAM hash matching to smaller services;
Response:
ii)   Please provide the underlying arguments and evidence that support your views.
Response:
iii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Question 23:
i)   Ability of services in scope of the CSAM hash matching measure to access hash databases/services, with respect to access criteria or requirements set by database and/or hash matching service providers;
Response:
ii)   Please provide the underlying arguments and evidence that support your views.
Response:
iii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Question 24:
i)   Costs of applying our CSAM URL detection measure to smaller services, and the effectiveness of fuzzy matching for CSAM URL detection;;
Response:
ii)   Please provide the underlying arguments and evidence that support your views.
Response:
iii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Question 25:
i)   Costs of applying our articles for use in frauds (standard keyword detection) measure, including for smaller services;
Response:
ii)   Please provide the underlying arguments and evidence that support your views.
Response:

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

Question 26:		
i)	An effective application of hash matching and/or URL detection for terrorism content, including how such measures could address concerns around 'context' and freedom of expression, and any information you have on the costs and efficacy of applying hash matching and URL detection for terrorism content to a range of services.	
Response:		
ii)	Please provide the underlying arguments and evidence that support your views.	
Response:		
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)	
Response:		

Automated content moderation (Search)

Question 27:		
i)	Do you agree with our proposals?	
Response: Please see above response to question 20		
ii)	Please provide the underlying arguments and evidence that support your views.	
Response: Please see above response to question 20		
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)	
Response:		

User reporting and complaints (U2U and search)

Question 28:		
i)	Do you agree with our proposals?	
Response:		
ii)	Please provide the underlying arguments and evidence that support your views.	
Response:		
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)	
Response:		

Terms of service and Publicly Available Statements

Question 29:
i) Do you agree with our proposals?
Response:
ii) Please provide the underlying arguments and evidence that support your views.
Response:
iii) Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Question 30:
i) Do you have any evidence, in particular on the use of prompts, to guide further work in this area?
Response:
ii) Please provide the underlying arguments and evidence that support your views.
Response:
iii) Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Default settings and user support for child users (U2U)

Question 31:
i) Do you agree with our proposals?
Response:
ii) Please provide the underlying arguments and evidence that support your views.
Response:
iii) Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Question 32:
i) Are there functionalities outside of the ones listed in our proposals, that should explicitly inform users around changing default settings?
Response:
ii) Is this response confidential? (if yes, please specify which part(s) are confidential)

Response:

Question 33:
i)   Are there other points within the user journey where under 18s should be informed of the risk of illegal content?
Response:
ii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Recommender system testing (U2U)

Question 34:
i)   Do you agree with our proposals?
Response:
ii)   Please provide the underlying arguments and evidence that support your views.
Response:
iii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Question 35:
i)   What evaluation methods might be suitable for smaller services that do not have the capacity to perform on-platform testing?
Response:
ii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

We are aware of design features and parameters that can be used in recommender system to minimise the distribution of illegal content, e.g. ensuring content/network balance and low/neutral weightings on content labelled as sensitive.

Question 36:
i)   Are you aware of any other design parameters and choices that are proven to improve user safety?
Response:
ii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Enhanced user control (U2U)

Question 37:	
i)	Do you agree with our proposals?
Response:	
ii)	Please provide the underlying arguments and evidence that support your views.
Response:	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:	

Question 38:	
i)	Do you think the first two proposed measures should include requirements for how these controls are made known to users?
Response:	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:	

Question 39:	
i)	Do you think there are situations where the labelling of accounts through voluntary verification schemes has particular value or risks?
Response:	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:	

User access to services (U2U)

Question 40:	
i)	Do you agree with our proposals?
Response:	
ii)	Please provide the underlying arguments and evidence that support your views.
Response:	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:	

Do you have any supporting information and evidence to inform any recommendations we may make on blocking sharers of CSAM content? Specifically:

Question 41:	
i)	What are the options available to block and prevent a user from returning to a service (e.g. blocking by username, email or IP address, or a combination of factors)?
Response:	
ii)	What are the advantages and disadvantages of the different options, including any potential impact on other users?
Response:	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:	

Question 42:	
i)	How long should a user be blocked for sharing known CSAM, and should the period vary depending on the nature of the offence committed?
Response: We believe that anybody sharing known CSAM should be permanently blocked from the service. There should be a zero-tolerance approach for sharing this content due to the significant harm caused to victims.	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:	

There is a risk that lawful content is erroneously classified as CSAM by automated systems, which may impact on the rights of law-abiding users.

Question 43:	
i)	What steps can services take to manage this risk? For example, are there alternative options to immediate blocking (such as a strikes system) that might help mitigate some of the risks and impacts on user rights?
Response:	
ii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:	

Service design and user support (Search)

Question 44:
i)   Do you agree with our proposals?
Response:
ii)   Please provide the underlying arguments and evidence that support your views.
Response:
iii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Cumulative Assessment

Question 45:
i)   Do you agree that the overall burden of our measures on low risk small and micro businesses is proportionate?
Response:
ii)   Please provide the underlying arguments and evidence that support your views.
Response:
iii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Question 46:
i)   Do you agree that the overall burden is proportionate for those small and micro businesses that find they have significant risks of illegal content and for whom we propose to recommend more measures?
Response:
ii)   Please provide the underlying arguments and evidence that support your views.
Response:
iii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Question 47:
i)   We are applying more measures to large services. Do you agree that the overall burden on large services proportionate?
Response:

ii) Please provide the underlying arguments and evidence that support your views.	
Response:	
iii) Is this response confidential? (if yes, please specify which part(s) are confidential)	
Response:	

Statutory Tests

Question 48:
i)   Do you agree that Ofcom's proposed recommendations for the Codes are appropriate in the light of the matters to which Ofcom must have regard?
Response:
ii)   Please provide the underlying arguments and evidence that support your views.
Response:
iii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Volume 5: How to judge whether content is illegal or not?

The Illegal Content Judgements Guidance (ICJG)

Question 49:
i)   Do you agree with our proposals, including the detail of the drafting?
Response:
ii)   What are the underlying arguments and evidence that inform your view?
Response:
iii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Question 50:
i)   Do you consider the guidance to be sufficiently accessible, particularly for services with limited access to legal expertise?
Response:
ii)   Please provide the underlying arguments and evidence that support your views.
Response:
iii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Question 51:
i)   What do you think of our assessment of what information is reasonably available and relevant to illegal content judgements?
Response:
ii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:

Volume 6: Information gathering and enforcement powers, and approach to supervision.

Information powers

Question 52:	
i)	Do you have any comments on our proposed approach to information gathering powers under the Online Safety Act?
Response:	
ii)	Please provide the underlying arguments and evidence that support your views.
Response:	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:	

Enforcement powers

Question 53:	
i)	Do you have any comments on our draft Online Safety Enforcement Guidance?
Response:	
ii)	Please provide the underlying arguments and evidence that support your views.
Response:	
iii)	Is this response confidential? (if yes, please specify which part(s) are confidential)
Response:	

Annex 13: Impact Assessments

Question 54:
i)   Do you agree that our proposals as set out in Chapter 16 (reporting and complaints), and Chapter 10 and Annex 6 (record keeping) are likely to have positive, or more positive impacts on opportunities to use Welsh and treating Welsh no less favourably than English?
Response:
ii)   If you disagree, please explain why, including how you consider these proposals could be revised to have positive effects or more positive effects, or no adverse effects or fewer adverse effects on opportunities to use Welsh and treating Welsh no less favourably than English.
Response:
iii)   Is this response confidential? (if yes, please specify which part(s) are confidential)
Response: