Your response

Question 1: To assist us in categorising responses, please provide a description of your organisation, service or interest in protection of children online.

Is this a confidential response? (select as appropriate)

No

Glitch is a UK charity (no. 1187714) that exists to end online abuse and to increase digital citizenship across all online users. We believe that our online community is as real as our offline one, and that everyone should work together to make it a better place. We work to promote good digital citizenship and address online harms such as online abuse, online hate speech and information disorders, and have developed bespoke training programmes covering Digital Citizenship, Online Active Bystanders and Digital Self Care and Self Defence. As part of this, we have delivered training to women in public life in addition to workshops in the classroom of schools, prior to 2020.

We are submitting evidence to Ofcom's inquiry because we believe that the Online Safety Bill regime has the potential to make a significant difference to the prevalence of online abuse experienced by internet users in the UK. However, for it to appropriately serve those disproportionately affected by online abuse – girls and women, and especially Black girls and women, and racialised and minoritised people – Glitch believes that the implementation of the proposed Online Safety Act will need to reflect the experiences of Black girls and women and other marginalised communities subjected to high levels of online abuse.

Recommendation: Violence Against Women and Girls Code of Practice

We believe that the Online Safety Bill needs to explicitly prioritise holding tech companies accountable for the gendered nature of online violence against women and therefore strongly recommend that Ofcom develop a specific violence against women and girls Code of Practice. Glitch has worked in partnership with the End Violence Against Women Coalition, Refuge, Carnegie, NSPCC, 5Rights and with experts Prof. Clare McGlynn and Prof. Lorna Wood on a model Violence Against Women and Girls Code of Practice, which we recommend Ofcom uses when writing their own VAWG Code of Practice, therefore embedding the deep knowledge of the experts working in this area into this vital piece of work for Ofcom.

Question 2: Can you identify factors which might indicate that a service is likely to attract child users?

Is this a confidential response? (select as appropriate)

No

It should be assumed that all services that are for social, leisure and entertainment purposes will likely attract child users, whether or not they are the intended audience of the content. Search, educational and information platforms etc will attract child users. Unless a service is designed for a specific purpose that is unlikely to be of interest to a child, for example an adult banking app, services by default should assume that under 18s can and will access their platforms, app and services, and mitigate against risks accordingly.

Question 3: What information do services have about the age of users on different platforms (including children)?

Is this a confidential response? (select as appropriate)

No

Facebook Whistleblower, Frances Haugen, has spoken many times about the high accuracy of Facebook's (now Meta) analysis of the age of its users, particularly in the context of children users. For example, Haugen gave evidence in the UK Parliament's Joint Committee on the draft Online Safety Bill on Monday 25 8 2021:

"Facebook has systems for estimating the age of any user. Within a year or two of them turning 13, enough of their actual age mates have now joined so that it can estimate accurately the real age of that person. In Facebook, you have to publish the protocols—how it does that—and the results going back a couple of years, and say how many 10 year-olds and how many 12 year-olds were on the platform one, two, three, four years ago. It knows this data today and it is not disclosing it to the public. That would be a forcing function to make it do better detection of young people on the platform."

It is concerning then that this information is only within the public domain from this specific company because of information brought to light by a whistleblower. and speaks to why regulation under a named regulator are so overdue.

This point from Haugen of course relates to just one parent company amongst many services within the scope of the Online Safety Bill but should be used as a case study when considering the information that platforms have or have not publicly disclosed in the past in order to increase transparency in the future under this regime.

Question 4: How can services ensure that children cannot access a service, or a part of it?

Is this a confidential response? (select as appropriate)

No

As outlined in question 3, platforms first need to be transparent about the technology they already have to estimate the ages of their users, and for that information to then support the deliberate attempts to ensure children cannot access a service or part of a service that is deemed to be for adult-only users. This is particularly pertinent considering the huge differences in the ways that the Online Safety Bill will be applied to children's online safety in comparison to the much less stringent or systemic application of the regime to adult users. Glitch's recommendation for Ofcom to write a VAWG Code of Practice aims to close some of this gap and would benefit children as well as adult users.

Question 6: Can you provide any evidence relating to the presence of content that is harmful to children on user-to-user and search services?

Is this a confidential response? (select as appropriate)

No

In the current online climate, all harmful content that can be seen publicly or that is private but could be viewed by a child on an account (either their own account or an adult, such as a parents' or guardian's etc) that is on user-to-user and search services is harmful to children. This may be true whether or not the child could be considered the intended audience or regardless of whether the company believes that children routinely use their service.

For example, harmful gender stereotyping and racism online is harmful to society as a whole, and when harmful content violates a platform's Terms of Service yet when reported is not removed, this gives all users, including children, the impression that such behaviours is permissible on the platform, and therefore acceptable. This is particularly pertinent in an environment where good digital citizenship education and media literacy for both children and adults is not widely available.

Girlguiding, EVAW and Glitch:

In more specific terms, Girlguiding's Girls' Attitudes Survey 2022 and subsequent research into on online harms highlights the presence of gendered harmful content experienced by girls and young women:

- Online harm comes in the form of sexist comments (35%), cyberflashing (22%), sexual harassment (20%), catfishing (20%), pressure to share nude pictures (16%) and cyberstalking (13%)

Question 6: Can you provide any evidence relating to the presence of content that is harmful to children on user-to-user and search services?

- 94% said they experienced negative emotions as a result with 76% saying it made them feel anxious, angry, scared, depressed or less confident in themselves
- 93% of girls and young women said there should be laws to protect against online abuse and 67% don't think the government is doing enough to stop online violence
- Only 15% think that social media is a safe place for them
- Research from Girlguiding showed that over three-quarters (79%) of young women have experienced online harm in 2021.

The End Violence Against Women Coalition (EVAW) have also recently launched a short film 'About Time' that focuses on sexual harassment at school, including online harassment, and published related survey results with young people that found:

- 80% of girls think schools need to do more to support young people's sex and relationships education, and to tackle sexual harassment in school
- Nearly three-quarters (72%) of young women say sexist behaviour makes them feel uncomfortable
- 62% of young women say comments about their body or uniform have made them feel uncomfortable
- 30% of young women don't feel safe from sexual harassment in school
- Almost a third of girls (32%) think schools wouldn't take reports of sexual harassment seriously
- 58% think racism is a problem at their school and 40% of those who have witnessed sexual name calling (and 46% of Black girls) have heard it reference race
- 60% think homophobia is a problem at their school, and 55% of those who have witnessed sexual name calling have heard it reference sexuality
- 1 in 4 girls have shared a sexual image of themselves (24%) and of those, a quarter (24%) said they felt pressured into it, and almost a third (31%) initially wanted to but later regretted it.
- Almost 1 in 4 (24%) girls in mixed sex schools say they have been the subject of unwanted sexual touching at school.

Ofsted research also supports the fact that these behaviours are happening in schools both online and offline.

Anecdotally, Glitch knows through work conducted in schools prior to the pandemic, that much of the discussions around online abuse related to harmful content in user-to-user and search services.

In Parliament:

In addition, in 2023, Alex Davies Jones MP, Baroness Merron and Lord Knight of Weymouth alongside others in parliament, have discussed the dangerous and emerging

Question 6: Can you provide any evidence relating to the presence of content that is harmful to children on user-to-user and search services?

trends around male 'influencers' such as Andrew Tate, (for example, the Lords' 2nd Reading debate on the Online Safety Bill, in the debate in the Common on 19 January 2022 debating Misogyny in Schools, and the debate Violence against Women and Girls: Plymouth on 25 January 2023).

In addition, Glitch amongst others, wrote to the Prime Minister raising concerns around the implications of misogynistic hate and so called influencers like Andrew Tate, recommending that a violence against women and girls Code of Practice be included in the Online Safety Bill.

Question 7: Can you provide any evidence relating to the impact on children from accessing content that is harmful to them?

Is this a confidential response? (select as appropriate)

No

Girlguiding's 2022 Attitudes Survey found that girls and young women don't feel safe in their daily lives, which included online, in public, when they were at school or on their own. 36% of the girls and young women surveyed (age 11-21) agreed that 'the abuse that high profile women get online' puts them off 'certain jobs (like politics). The survey found a regional discretion (41% in N England compared with 34% in London and the South). Worryingly, the rate was might higher for disabled girls and young women (49%) compared to survey respondents who aren't disabled. Young women of colour (17-21 year olds) wer more likely to be put off than white girls (42% compared to 34%).

The survey also highlighted the importance of anonymous accounts increasing feelings of safety, which was especially the case with LGBTQ+ girls and young women (40%) compared with 27% non LGBTQ+. This needs to be considered when considering verification.

Though small snapshots, both examples show the need for intersectional analysis of these harms and a response that factors in protected and unique characteristics (that both reflects and goes further than those listed in the Equality Act 2010 e.g. including non-binary identities which are currently not recognised in UK law).

More statistics from Girlguiding's research is included in the answer to question 6.

Cyber Bullying:

Research into children and bullying, including cyberbullying highlights the inescapability of online abuse, such as cyberbullying. Home is no longer a safe place, for children or adults subjected to online abuse. One cannot simply leave the street or the school gates where

Question 7: Can you provide any evidence relating to the impact on children from accessing content that is harmful to them?

verbal insults may have taken place, as is possible when bullying/abuse is perpetrated solely offline.

“Cyberbullying is not limited to a specific location, such as a school or club, therefore, victims can be targeted in any place, and at any time. Unlike traditional bullying, the home no longer represents a safe place. Cyberbullying can follow a victim, invading all aspects of their personal life, and allowing them little opportunity to escape”
Anti-Bullying Alliance ‘Focus on Cyberbullying’ report p.3 (author Neil Tippett)

The speed that online abuse can circulate online is also a concern, adding to its impact on the person targeted with the abuse.

“Cyberbullying is capable of reaching a far broader audience than more traditional forms of bullying. Initial incidents of cyberbullying, such as posting an embarrassing photo or video, can spread throughout social networks, traversing school and personal boundaries, and increasing the chance that others will join in with the bullying. That cyberbullying can happen anywhere, and involve multiple, potentially anonymous perpetrators, has made it particularly difficult for schools to know when and how to respond to incidents of cyberbullying.”
Anti-Bullying Alliance ‘Focus on Cyberbullying’ report p.4 (author Neil Tippett)

The Facebook whistleblower Frances Haugen further discusses the way that the prevalence of online engagement, including negativity and online abuse and cyberbullying creates a different lived experience for teens compared to their parents:

“It used to be that it didn’t matter how bad school was, you could go home at the end of the day, and the vast majority of kids have good home lives, they got a break… it didn’t matter how badly you got bullied, you got a solid 16 hours to reset before you went back into the fray. And now that bullying of that harassment follows kids into their bedrooms. Like the last thing kids see before they fall asleep at night is someone being cruel to them. Or they wake up in the morning to some horrible slur about them or their personality. That really wears kids down and Facebook know that parents don’t have the right context about neurological development, they don’t have context on what is or isn’t effective for coaching kids on how to deal with these situations. And they give advice like why don’t you just turn off your phone, why don’t you just not use that. And the reality is that kids feel fear of being ostracized. Well that’s in Facebook’s docs if they don’t use the product. That kids say I know it doesn’t make me happy, I know I’m not having fun, I know I don’t want to use it, but I also feel like I can’t stop.

Question 7: Can you provide any evidence relating to the impact on children from accessing content that is harmful to them?

I think there is a real role for schools in helping to pull together high-quality information and make sure that every parent gets that information regularly on what is or isn't constructive in terms of how to coach and support kids. Because right now parents are trying really hard and because there's those gaps, those differences of lived experience, the parents just aren't being set up to succeed and the parents are struggling just as much as the kids are." Facebook Whistleblower Frances Haugen 'The Facebook Files: What's Next? Panel 1: The Activists' event hosted by Yale ISP, Thursday 7 October 2021.

Question 9: What are the exacerbating risk factors services do or should consider which may have an impact on the risk of harm to children in the UK?

Is this a confidential response? (select as appropriate)

No

Evidence on online harms frequently highlights the way in which intersecting discrimination exacerbates harm. It is essential that protected and unique characteristics are factored into services' plans and actions when risk assessing, mitigating and ending online harms on their platforms.

For example, women are 27 times more likely to be harassed online than men, with Black women 84% more likely to be targets of abusive tweets than white women and 60% more likely to receive problematic tweets.

As supported by adult research on intersecting impact of discrimination (e.g. Glitch and the End Violence Against Women's 2020 Ripple Effect report; Glitch's forthcoming research on misogynoir online (May 23)) and Girlguiding research (as listed above), intersecting discrimination and lived experience relating to protected and unique characteristics has a profound effect on users' experiences. Therefore, risk assessments must ensure that protected and unique characteristics are taken fully within context and into account. No one-size-fix all approach will bring change to the most affected users on platforms. In relation to Glitch's work, a gender-neutral approach that does not factor in racial discrimination cannot deliver deep and meaningful change for Black women and girls on platforms and services.

Question 11: What can providers of online services do to enhance the clarity and accessibility of terms of service and public policy statements for children (including children of different ages)?

Is this a confidential response? (select as appropriate)

No

Tech companies should ensure they:

- use age-appropriate language
- use engaging presentation of the information
- increase user trust in their companies – e.g. through enforcement of their terms of services, and appropriate and effective reporting mechanisms that are trauma-informed.

User habits in relation to interactions to terms of service and public policy statements for children should be analysed to ensure meaningful interactions between children and platforms.

Question 13: What can providers of online services do to enhance children's accessibility and awareness of reporting and complaints mechanisms?

Is this a confidential response? (select as appropriate)

No

Online services need to build trust through action by delivering on the standards they have set out themselves in their terms of service. Their reporting and complaints mechanisms should lay out expectations for users and meet those expectations, as broken trust leads to non-reporting and the perpetuation of harm to the original user and others. If platforms fail to demonstrate good reporting and complaints mechanisms children on platforms will continue to not use them as seems to widely be the case currently, and this lack of trust will disseminate amongst children both through word-of-mouth as well as because of personal user experience of being let down by the existing reporting and complaints mechanism. One issue that platforms face now is that they will need to regain trust that has already been lost by demonstrate much needed improvements to these mechanisms.

The regulator should ensure that services' reporting and complaints mechanisms meet a high standard and are being delivered - level of trust in platforms reporting and complaints mechanisms should be monitored by Ofcom.

Question 15: What actions do or should services take in response to reports or complaints about online content harmful to children (including complaints from children)?

Is this a confidential response? (select as appropriate)

[Please select]

Services should consistently enforce the actions that they have outlined in their terms of service and policies. Terms of service should meet a mandated minimum standard and be robust. *See answer to question 13 for more on rebuilding lost trust.*

Question 16: What functionalities or features currently exist that are designed to prevent or mitigate the risk or impact of content that is harmful to children? A1.21 in the call for evidence provides some examples of functionalities.

Is this a confidential response? (select as appropriate)

No

Different platforms have different functionalities which can have negative and positive impacts in relation to safety. For example, 'disappearing messages' in WhatsApp may increase the chances of losing evidence of harm, while it may also increase the chances of messages being used for harm in the future.

While 'snapshots' of messages can mitigate this on one platform such as WhatsApp, the same action can have quite different results on other apps, such as BeReal - where users are notified that a named user has taken a screenshot of their image. In this case, the alert may alert the user who posted the content to potential harm pinpointed to a particular, named user who received it.

The benefit of healthy and fair competition within the tech market can allow more user-choice when it comes to selecting which functionalities are best suited to any individual user, including children.

There is of course, strong debate around encryption and children's safety in relation to CSEA, which is not within the remit of Glitch's submission.

Question 17: To what extent does or can a service adopt functionalities or features, designed to mitigate the risk or impact of content that is harmful to children on that service?

Is this a confidential response? (select as appropriate)

No

Features and functionalities that enhance user empowerment, personal curation and that respond to a changing environment of harm are one important factor in mitigating the risks or impact of content that is harmful to children.

As Glitch argues in relation to harm to adults in the Online Safety Bill, these features and measures are one element of a wide range of possible mitigants to this harm, and do not serve to solve the issues when applied to harmful online content, without further provision.

For this reason, Glitch campaigns for increased tech company accountability in relation to platforms and services taking a systemic approach to mitigating and ending online harms, which includes many elements, as outline in our jointly written* Violence Against Women and Girls Code of Practice, which is highly relevant to the Online Safety Bill's response to harms to children and was written in partnership with children's charities and others (listed below).

The areas covered within the Code, which goes beyond services adopting functionalities and features designed to mitigate risk includes:

1) Responsibility, risk assessment, mitigation and remediation
2) Safety by design
3) Access to online service, terms of service and content creation
4) Discovery and navigation
5) User response, user tools
6) Moderation
7) Transparency
8) Victim support and mediation
9) Safety testing
10) Supply Chain Issues
11) Enforcement of criminal law
12) Education and training
13) Vigilance over time

An amendment on the inclusion of a Violence Against Women and Girls Code of Practice has been tabled by Baroness Morgan in the House of Lords.

**in partnership with the End Violence Against Women Coalition, Refuge, Carnegie, NSPCC, 5Rights and with experts Prof. Clare McGlynn and Prof. Lorna Wood.*

Question 18: How can services support the safety and wellbeing of UK child users as regards to content that is harmful to them?

Is this a confidential response? (select as appropriate)

No

The Children's Code of Conduct for services should be written to ensure that services support child safety and wellbeing in a way that reflects their lived experience on platforms, including based on their protected and unique characteristics.

For example, a gendered perspective should be applied to child-user experiences, expertise of children's charities and online gender based violence organisations and researchers should inform this work.

A mandated violence against women and girls Code of Practice would support the systemic approach to this harmful content by placing systematic requirements and understanding on platforms. The specifics mapped out in the model VAWG Code of Practice should also inform the Children's Code of Practice.

Question 23: What training and support is or should be provided to moderators?

Is this a confidential response? (select as appropriate)

No

Human moderators are vital to the process of moderation, adding local and cultural context, human sophistication and complex decision making to nuanced material. However, there is a huge emotional and psychological cost to this difficult work, which must be acknowledged, remunerated and the potential harms to human moderators mitigated against.

- Tech companies need to invest more in human moderation and ensure moderation considers local context, including (but not limited to) linguistic, social, cultural, historical, racial and gendered context.
- Human moderators should work in holistic environments which appropriately support their wellbeing, proportionate to the level of upsetting and harmful material they are moderating
- Diversity within teams is incredibly important, as is the training that is offered, for example training on recognising and responding appropriately to racism; online gender based violence; anti-transphobic content etc.
- Human moderators should be paid well in recognition of the heavy burden of a difficult job.
- Comprehensive training for moderators about online gender-based violence and different tactics of online abuse, and how abuse specifically targets women, Black

Question 23: What training and support is or should be provided to moderators?

and minoritised communities and users with intersecting identities is paramount - without this moderation risks being ineffective, inequitable and/or discriminatory

- Tech companies need to be transparent about their investment in and resourcing of content moderation
- It is essential that users understand when content has been moderated by human moderators and when it has been moderated by other means.
- Platforms should acknowledge internet biases in machine learning and AI systems and aim to eliminate biases through trusted partner interventions, for example through 'bias bounty challenges' and open access for researchers if these approaches are deemed to be effective by Ofcom.

Glitch's work with Social Finance for the World Wide Web Foundation report 'Strengthening Accountability for Online Gender-Based Violence – one year later' also looked at the important role of human moderators with regionally specific reflections. Learnings from these findings should be applied to the nature of online abuse across the UK, for example in relation to national context, dialects and languages (e.g. national languages across the four nations of the UK such as Cymraeg/Welsh, Irish Gaelic and Scottish Gaelic) as well as minority languages that are spoken in and across the UK (as highlighted through ONS data for example). A system of moderation that largely focuses on American English due to the market dominance of Silicon Valley tech companies will not functionally deliver high quality moderation within the UK - this regionality even within the English language was highlighted by Facebook whistleblower Frances Haugen when she gave evidence in the UK Parliament's Joint Committee on the draft Online Safety Bill on Monday 25.8.2021:

"I am deeply concerned about its underinvestment in non-English languages and how it misleads the public into thinking that it is supporting them. Facebook says things like, "We support 50 languages", when, in reality, most of those languages get a tiny fraction of the safety systems that English gets. Also, and I do not think this is widely known, UK English is sufficiently different that I would be unsurprised if the safety systems that it developed primarily for American English were underenforcing in the UK. Facebook should have to disclose dialectical differences."

Though of course relating to Meta, this evidence should be considered when considering content moderation models and the importance of human moderators with local context and connections across tech services.

More information on moderation is included in the VAWG Code of Practice.

Question 24: How do human moderators and automated systems work together, and what is their relative scale? How should services guard against automation bias?

Is this a confidential response? (select as appropriate)

No

By default, automated systems should be assumed to be biassed and tested accordingly by external researchers, in order to create better, more robust systems that are consistently improved to reduce automation bias. Awareness of the likelihood of biases and transparency around data sets is key to combating automation bias.

Question 28: Other than those covered above in this document (the call for evidence), are you aware of other measures available for mitigating the risk, and impact of, harm from content that is harmful to children?

Is this a confidential response? (select as appropriate)

No

Tech companies may state that their services are not the source of the societal harms that are perpetrated on their platforms, which may be an argument used to diminish the responsibility of companies' when it comes to the deliberate design choices and business models which, whether inadvertently or not, fail by allowing, amplifying harmful content. It is clear that tech companies have failed to self-regulate their platforms and services within the framing of their own terms of service and the time for legislated regulation is long overdue.

Online harms contribute to a wide public discourse which intersects with the way in which the media, government and politics, education, policing and criminal justice system all contribute to societal harms that manifest online.

Glitch is clear that a wide and far reaching public health approach to mitigate online abuse across the UK is needed. This approach should include anti-colonialist and anti-racism curricula in schools, strong systemic and regionally consistent approaches to ending violence against women and girls and a society-wide media literacy campaign which aims to create wider understanding of the issues and impacts of online harms and digital citizenship education for the adult and youth and children populations.

Understanding the nature and breadth of the issues is fundamental to finding meaningful and long lasting solutions to mitigating the risk and impact of harm from content that is harmful to children.

Question 28: Other than those covered above in this document (the call for evidence), are you aware of other measures available for mitigating the risk, and impact of, harm from content that is harmful to children?

It cannot be stated strongly enough that ‘legal but harmful’ content aimed at adults that is out of scope in the Online Safety Bill as it is currently drafted is harmful towards children and society as a whole. The two tier system of internet - as the bill proposes - that includes robust regulation when it comes to child-users that falls away when a young person reaches their 18th birthday is both dangerous and unfit for purpose. This is clear when comparing the lived experience of young people, who state that online harm, whether legal or otherwise, increases as they get older, and therefore increases throughout childhood into early adulthood. These trends can be seen in research including the previously mentioned Girlguiding Girls’ Attitudes Surveys.

Our children and young people deserve a regime and approach to online safety that listens to their voices, is rooted in their lived experiences and responds accordingly.