



**Pinterest Response to Consultation: Protecting children from harms online**  
**17 July 2024**

<b>Consultation Title</b>	Protecting children from harms online
<b>Organisation Name (if applicable)</b>	Pinterest

## **Introduction**

The following is Pinterest’s response to the consultation published on 8 May 2024 by the Office of Communications (Ofcom) on how online platforms should approach their new duties under the Online Safety Act (OSA) relating to content that is harmful to children, including the Draft Children’s Access Assessments Guidance, Draft Children’s Risk Assessment Guidance (collectively with previous, the “Guidance”), and Draft Protection of Children Code of Practice for User-to-User Services (the “Code”).<sup>1</sup> Pinterest welcomes the opportunity to respond to the consultation, as we share Ofcom’s goal of fostering a safe and positive online environment, particularly for more vulnerable users such as teens, and support its objective of creating guidance and codes of

---

<sup>1</sup> As the Draft Children’s Safety Code for Search Services is not applicable to Pinterest, it is not within the scope of this response.

practice that are proportionate and adaptable to the varied platforms that are subject to the OSA.

Pinterest is a visual discovery engine where people around the world go to get the inspiration to create a life they love. Browsing and saving visual ideas on Pinterest's service helps users imagine what their future could look like, and go from inspiration to reality.

Pinterest shows users visual recommendations called Pins. Pins are created by both individuals and businesses by uploading photos or videos or bookmarking content from the web, and providing a text caption. Users can save and organise these recommendations into collections called boards. Some of our most common use cases include food, beauty, home decor, and travel.

Pinterest has a longstanding commitment to making our platform a safe place for everyone, which is the right thing for the people who use Pinterest and aligns with our goals as a business. We view a positive, inspiring atmosphere as essential to our user experience, and accordingly, we have developed Community Guidelines that broadly prohibit various types of content that may be harmful to users of all ages, including, among others: pornography and sexually explicit content; content promoting suicide or self-harm; hateful or violent content; content promoting disordered eating; and health, climate, or civic misinformation. While all content that violates our Community Guidelines might be considered harmful to teens, there are also specific areas of our policies that focus on harms that disproportionately affect younger users, such as our prohibition of harassment and criticism, and of dangerous pranks or challenges. In other areas, content that might be permitted in other circumstances is prohibited when it involves minors, for example, depiction of alcohol use.<sup>2</sup>

We enforce these policies through automated tools, manual review, and hybrid approaches that combine elements of both, and we work with trusted third parties and the wider industry to enhance our work. In addition to global enforcement of our content policies, Pinterest utilises a number of features that are specifically designed to promote teen safety on our platform. For example, Pinterest offers a private profile as the default and only option for users under the age of 16, and the default option for users aged 16 and 17. Private accounts are undiscoverable on Pinterest search and search engines, and profiles, boards and Pins for these users will only be visible to followers approved by the user. Users with private profiles are able to connect with family and friends by sending a unique profile link, and all users have the ability to review and remove followers. Further, teens under the age of 16 can only send and receive messages from

---

<sup>2</sup> Pinterest's Community Guidelines can be accessed at: <https://policy.pinterest.com/community-guidelines>

mutual followers who they have invited to follow them through a unique profile link. Teens under the age of 18 can only receive messages from mutual followers, and can only receive message requests from users they follow. Pinterest also offers a Parental Passcode feature for users under the age of 18, which allows parents and caregivers to set up a 4-digit passcode to lock certain settings and have oversight of their teen's Pinterest account.<sup>3</sup>

These efforts are effective in providing a high level of safety and privacy for teen users on Pinterest, but we know that the work needed to achieve our goal of providing a safe and positive platform for our users will never end. We're committed to keeping Pinterest a positive corner of the internet. That is why we continually review and update our safety policies, tools, and technologies. Keeping our community safe is an industry-wide challenge as online trends and technological solutions continue to evolve, and we remain committed to exploring the best ways to achieve this goal. Indeed, as a founding signatory of The Inspired Internet Pledge, Pinterest has publicly pledged to centre its commitment to making the internet a healthier place through its product, thought leadership, policies, and strategic giving.<sup>4</sup>

Our commitment to teen safety comes from the top, as our Chief Executive Officer Bill Ready has been outspoken on the importance of making user well-being, in particular teen safety, central to our company's purpose. For example, earlier this year, shortly before the United States Senate heard testimony from technology executives on their efforts to protect children online, Mr. Ready published an op-ed stressing the importance of placing artificial intelligence at the centre of improving youth online safety and wellbeing, as well as for platforms to add special protections for teens. In doing so, he urged our entire sector to take accountability for this issue and make it part of our business model: "It's time for industry leaders to accept responsibility and build online platforms centred on positive well-being outcomes. [...] A successful social media business model that is focused on positivity is possible. I know this because we are building it here and seeing results."<sup>5</sup>

We expect that we will already be compliant with significant segments of the Code thanks to our stringent content safety policies, robust content moderation systems, and

---

<sup>3</sup> More information on these features can be found in Pinterest's Teen Safety hub, as well as our Help Centre:

<https://business.pinterest.com/teen-safety-on-pinterest/>

<https://help.pinterest.com/article/teen-safety-options/>

<https://help.pinterest.com/article/resources-for-parents-and-caregivers-of-teens/>

<sup>4</sup> For more information on The Inspired Internet Pledge, please see:

<https://inspiredinternet.org/signatory/pinterest/>

<sup>5</sup> The entire op-ed is available in Pinterest's newsroom:

<https://newsroom.pinterest.com/news/pinterest-ceo-big-tech-must-reject-big-tobaccos-model/>

strong measures for the protection of minors. However, as detailed herein, we believe certain sections of the Guidance and Code either require clarification or impose requirements that are not effective, practicable, or proportionate in the context of Pinterest's functionality and common use cases.

This response is divided into two sections, with the first containing general comments that are applicable throughout Ofcom's proposals, and the second consisting of responses to specific questions posed by the consultation.

## **General Comments**

Pinterest's general comments address (1) the importance of proportionality in drafting and applying the Guidance and Code; (2) whether and how platforms may implement measures other than those included in the Guidance and Code; and (3) the proposed implementation timelines.

### **The Importance of Proportionality in Guidance and Codes of Practice**

Pinterest is encouraged by Ofcom's stated intention that the Guidance and Code be implemented in a manner that is proportionate to platforms' differing risk profiles and operational capacities. In order to preserve competition in the digital sector and allow platforms to implement measures that are suitable and effective for their own platforms, we agree that it is crucial for the Guidance and Code to be interpreted and applied in a manner that is tailored to the OSA's objectives and ensures compliance is practicable for the various platforms in the law's scope. We also welcome the numerous areas of alignment between the Draft Protection of Children Code and the Draft Illegal Content Code, which helps to streamline platforms' compliance efforts.

The proportionality of the Guidance and Code's measures may be assessed by taking account of a range of factors, including the size of the platform, its functionalities and use cases, the risk to users, existing mitigation measures, relevant industry standards, users' rights, and the platform's resources compared to the cost and effectiveness of the proposed measures. In view of these factors, it's clear that platforms are exceedingly diverse, and accordingly the Guidance and Code should allow flexibility for platforms to implement only the safety measures that are practical and effective in light of their unique designs and risk profiles, and to do so in a manner that is suited to the features and functionality of their platforms.

In order for small- and medium-sized platforms to continue to innovate and grow, it is important that these platforms not be required to take steps that are unnecessary to

improve their users' safety, considering the substantial aggregate costs of compliance with the Guidance and Code's measures. Even platforms that may be considered "large" due to their user base may have vastly different resources and capabilities than other "large" platforms due to differences in their business models or stages of development. As such, we recommend a nuanced approach that accounts for platforms' respective features, risk profiles, and resources, to ensure that Ofcom's regulatory framework is tailored to the relevant circumstances and avoids undue burden on platforms wherever possible.

### Implementation of Alternative Measures

We recognize that Ofcom has emphasised that the Guidance and Code consist of recommendations, and that platforms may comply by implementing alternative measures, as the ultimate goal of the online safety regime should be for platforms to implement mitigations that are suitable and effective for achieving their intended purpose. However, as detailed herein, some measures throughout the Guidance and Code appear to be highly prescriptive, specifying particular measures or technologies rather than setting an outcome to be achieved and allowing regulated services the flexibility to implement proportionate solutions that are tailored to their platforms.

While these measures are put forth as recommendations, it is not clear how platforms can demonstrate that their alternative measures are achieving a similar (or better) outcome, or how Ofcom would assess whether those alternative measures are sufficient to comply with the Guidance and Code. Ofcom's recognition of alternative approaches to compliance is welcome, and while we appreciate Ofcom's desire to maintain flexibility in assessing them, guidance on the criteria to be applied to alternative measures would help platforms confidently plan for compliance with the Guidance and Code and streamline the implementation process across platforms. This would also prevent the Guidance and Code's "recommended" measures from effectively becoming required, as platforms would otherwise feel pressure to default to them to avoid risk of enforcement. Certain of these recommended measures will not be effective or practicable for all entities in scope of the OSA, and in fact may be suitable only for well-resourced incumbent platforms.

### Implementation Period

The OSA states that platforms must complete their risk assessments within three months of Ofcom publishing final statements. While the legislation provides the regulator with flexibility to negotiate alternative timelines, to date Ofcom has continued to propose three-month timetables to carry out the Children Access and Children's Risk Assessments, as well as to put required mitigation measures into place.

In the interests of effectuating principles of proportionality, as well as giving legal certainty to platform operators, we urge Ofcom to exercise its discretion when it comes to establishing time periods for implementation. Even with the foresight provided by the consultation process, a timetable of three months from when requirements are finalised is extremely challenging, particularly for implementing remedial tools and processes which require significant planning and resource allocation by numerous internal teams to effectively design, develop, test, and launch the required changes. This is especially so for small- and medium-size platforms, which need to carefully allocate limited resources to comply with new requirements in this and other jurisdictions.

As such, we respectfully request that Ofcom give consideration to extending implementation periods in response to feedback from platforms, particularly considering the complexity of the requirements and the resources of the platform in question, or that it at least refrain from resorting directly to an enforcement posture should platforms demonstrate in good faith that additional time is needed to come into compliance with the Code.

## **Responses to Consultation Questions**

### **Identifying the services children are using / Children's Access Assessments**

*Do you agree with our proposals in relation to children's access assessments, in particular the aspects below? Please provide evidence to support your view.*

*2. Our proposed approach to the child user condition, including our proposed interpretation of "significant number of users who are children" and the factors that service providers consider in assessing whether the child user condition is met?*

The minimum required age to register for Pinterest is 13 years old (or older in countries with a higher age of digital consent), so there is no question that users under 18 access our service. We are glad to see the proportionate approach reflected in Ofcom's proposal to treat access by a "significant number" of children as a further benchmark, but considering that the existing proposal does not define a "significant" population of users, we would appreciate any additional guidance on how this benchmark should be assessed.

## Draft Guidance on Content Harmful to Children

*12. Do you agree with our proposed approach, including the level of specificity of examples given and the proposal to include contextual information for services to consider?*

We acknowledge Ofcom's thorough approach in providing detailed definitions, examples, and counterexamples of the various types of harmful content in scope of the Code. It is undoubtedly challenging, though, to categorise the myriad forms of content that may be found online, and unsurprisingly there are several areas where the proposed guidance for defining harmful content is subjective and would thus be difficult to apply consistently at scale. For example:

- The Eating Disorder Key Terms definition for "Promote" states that "Promotion could be unintentional or inadvertent. This can include content which glamourises, glorifies, romanticises, or normalises eating disorders." The broad scope of this definition suggests that promotion should be defined by the viewer's subjective reaction to content rather than through objective criteria, which may lead to inconsistent application.
- One of the content examples under the Eating Disorder guidance is: "Content which provides instructions for weight loss or fitness which is extreme and/or excessive and/or dangerous, for example posts promoting excessive exercise despite injury or fatigue." What constitutes "extreme" fitness in this context is highly subjective.
- The guidance on Pornography suggests that there is a difference between content that is pornographic and content that is sexually suggestive, but there is no definition of "sexually suggestive" or counter-examples illustrating the difference. The contextual guidance provided is very subjective.

We appreciate Ofcom's efforts to promote online safety by defining harmful content broadly, but also recognize that subjective concepts are exceedingly difficult to apply to the varied content hosted by online platforms, as reasonable people could reach different conclusions on whether a given item of content is harmful. We are concerned, then, that platforms implementing the Code may feel pressure to over-enforce against safe content, or potentially be subject to enforcement despite their efforts to moderate content in good faith.

In light of the inherent difficulties in defining harmful content, we believe it is advisable for these definitions to be applied flexibly while assessing platforms' content policies, and to the extent Ofcom has concerns over the scope of a platform's policies, that this area should be treated as ripe for further engagement, as opposed to entering an enforcement posture.

## Children's Risk Assessment Guidance and Children's Risk Profiles'

*17. What do you think about our proposals in relation to the Children's Risk Assessment Guidance?*

*-and-*

*19. Do you think the four-step risk assessment process and the Children's Risk Profiles are useful models to help services understand the risks that their services pose to children and comply with their child risk assessment obligations under the Act?*

We believe that the four-step risk assessment process and the Children's Risk Profiles are useful models which generally align with best practise. We advise, however, that it will be onerous for platforms to implement the assessments as outlined in the proposed guidance, requiring a significant amount of cross-functional resources, including time from staff who are dedicated to ensuring user safety.

In order to alleviate this burden, we propose that Ofcom continue to demonstrate flexibility and proportionality in the implementation of risk assessments, allowing platforms to assess risk in a manner that is suited to their unique services.

Ofcom can also assist platforms' compliance by aligning with international requirements to the extent possible. As platforms are increasingly being required to produce risk assessments in jurisdictions around the world, we would be eager to partner with Ofcom to find opportunities to streamline these efforts, both to conserve costs and also to simplify the monitoring work done by Ofcom.

It may also be helpful to ultimately consolidate the illegal harms and child safety risk assessments into a single process. We understand from prior comments by Ofcom that this is not feasible for the initial risk assessments due to the phased approach to implementing the OSA, but in future years a consolidated approach would be less burdensome while also providing a clearer picture of how platforms assess and mitigate risk, as these efforts are often interconnected across categories of harm.

## What should services do to mitigate the risk of online harms / Our proposals for the Children's Safety Codes

*22. Do you agree with our proposed package of measures for the first Children's Safety Codes?*

Pinterest generally agrees with the proposed package of measures for the Code, as it reflects principles of effective content moderation, strong governance and accountability,



and robust choice and support for children that are aligned with our current practices. As mentioned in our general comments, however, we welcome continued engagement with Ofcom on implementing the Code's objectives in a manner that is tailored to Pinterest's functionality and risk profile, rather than imposing prescriptive requirements.

*23. Do you currently employ measures or have additional evidence in the areas we have set out for future consideration? If so, please provide evidence of the impact, effectiveness and cost of such measures, including any results from trialling or testing of measures?*

One area set out by Ofcom for future consideration is the use of automated detection tools. Pinterest agrees with the utility of these tools, and indeed we currently implement them to identify and action some types of violative content. We urge caution, however, when it comes to mandating their use in future codes, as such a requirement may be disproportionate.

Automated content moderation can be a blunt instrument, actioning content without fully accounting for context or nuance. Particularly with regards to categories of harmful content that do not lend themselves to identifiable imagery or language,<sup>6</sup> automated systems may be prone to capturing a large amount of safe content, while also failing to capture actual harmful content. This creates risks of over-enforcement against unobjectionable content, as well as potentially stifling free expression online, without completely eliminating harmful content.

Automated content moderation systems are also costly to develop and maintain, and even after they are implemented, human agents are required to support them by reviewing edge cases captured by the system, handling user appeals, and, where applicable, labelling content to continue to train machine learning models. These tasks would be especially burdensome if a future code is overly prescriptive and requires a system to be calibrated in such a way that it captures a large number of false positives. Such operational burdens may only be practicable for the largest and most well-resourced platforms.

Rather than broadly requiring the use of such systems, then, they should be encouraged only for platforms that are at high risk for harm categories where automated content moderation can be effectively deployed. Furthermore, platforms should be granted the flexibility to iterate their systems in order to find the right balance between detecting harmful content and avoiding over-enforcement.

---

<sup>6</sup> For example, see Pinterest's Response, dated 23 February 2024, to Consultation: Protecting people from illegal harms online ("Illegal Harms Consultation"), at Question 16, Measure 4I: Keyword detection regarding articles for use in frauds.

## Developing the Children's Safety Codes: Our framework

*27. Do you agree that most measures should apply to services that are either large services or smaller services that present a medium or high level of risk to children?  
-and-*

*29. Do you agree with our definition of 'multi-risk' and with how we apply this in our recommendations?*

As mentioned in the prior consultation on illegal harms, Pinterest appreciates the approach taken by Ofcom in tiering its guidance and codes of conduct according to platform size and risk level, but in practice its proposed risk thresholds will place a disproportionate burden on many platforms.<sup>7</sup>

We believe that categorising a service as “large” based on number of users does not by itself reflect its risk profile or level of resources. Despite having a large number of users, a platform may actually have a low risk profile once its features and functionalities, common use cases, and existing risk mitigation measures are taken into consideration. A platform with a large user base may also have vastly different resources and capabilities than other “large” platforms due to differences in its business model, market conditions, or its stage of development as a company. As such, imposing requirements on platforms merely due to their size would be disproportionate and give the Code a much more prescriptive character, instead of adapting and responding appropriately to the actual risks posed by a particular service.

In addition, as stated in our comment on the definition of “multi-risk” in the consultation on illegal harms, as a general principle we do not agree that being at medium risk of at least two harms necessarily entails a significant overall risk, particularly when circumstances such as a platform’s design, use cases, and existing mitigation measures are taken into consideration.

This is particularly true in the context of this consultation, in which content harmful to children is defined very broadly. While we believe this is well-intentioned, it is also likely to result in a large number of platforms being found to carry heightened risks of harmful content. Simply tallying these inherent risks is not adequate to assess the actual risk posed to users, however. To the extent a platform may carry some risk if left untreated, it is important to consider what controls and safeguards it already has in place to mitigate that risk, such as content safety guidelines and content moderation systems, as these will result in a smaller, residual risk for each type of harmful content. Failing to

---

<sup>7</sup> See Pinterest Response to Illegal Harms Consultation, at Question 15.

account for these controls would effectively eliminate risk profile as a factor in determining which obligations apply to which platforms. It could also have the unintended effect of discouraging industry-led innovation in new safety techniques and processes, by not acknowledging that platforms either currently have, or could in the future, develop and implement techniques which go beyond, or are more effective than, Ofcom's proposed remedies. This is particularly the case where platforms develop or implement new methods which are more relevant to their platforms' use cases and purpose.

We believe that despite its intention to differentiate between platforms, Ofcom's proposal will foreseeably lead to most platforms being categorised as multi-risk. This would likely lead to the most onerous measures being applied to platforms whose risk profiles do not merit such treatment, thus negating Ofcom's efforts towards proportionality. As an alternative, we would welcome a more holistic, nuanced approach that accounts for the role of platforms' design and risk mitigation measures. This would more accurately reflect the true risk profiles of regulated platforms, and help foster a safe and positive online experience while avoiding undue burden on platforms that already provide a low-risk service to teen users.

#### Age assurance measures

*31. Do you agree with our proposal to recommend the use of highly effective age assurance to support Measures AA1-6?*

- a) Are there any cases in which HEAA may not be appropriate and proportionate?*
- b) In this case, are there alternative approaches to age assurance which would be better suited?*

Pinterest supports a risk-based approach to age verification that balances the protection of minors with the principle of data minimisation, as well as the relative risk of harm to minors on the platform. Under this approach, platforms should consider the type of service being offered, the risk of harm to minors, and data minimisation obligations to ensure they collect only the data necessary to reach a proportionate degree of certainty about the age of its users.

In pursuing its own risk-based approach to age verification, Pinterest considers factors including its use case as a visual discovery engine for helping users find ideas like recipes and home and style inspiration, as well as the scope of its Community Guidelines. In order to provide a safe and inspired experience for both teens and adults, Pinterest has developed strong Community Guidelines with our youngest users in mind and applies them to all users and content on Pinterest. For example, unlike some other

platforms, Pinterest does not age-gate sexualized content because sexualized content is prohibited on Pinterest. As Pinterest's content policies and enforcement are designed to provide a safe experience for all users, we are concerned that requiring additional age verification measures from individual services would require collection of sensitive user information without any corresponding improvement in user safety.

We further note that not all services are able to effectively estimate age based on use of their service. Where a service does not encourage the sharing of personal data, any such speculation as to a user's age is likely to result in false positives, as well as unnecessary processing of personal identity document information in connection with those false positives. For example, an adult user may interact with content for children as they are looking for gift ideas for their children, planning their child's birthday party, or lesson-planning for a school curriculum. A predictive mechanism could potentially flag those users, inaccurately, as being underage. The user would then need to provide sensitive information to prove their age, such as an identity document, in order to continue using the service.

In light of the potential risk to users' privacy and security from increased data collection, we believe that the final Code should ensure flexibility for platforms to only implement additional age verification measures which are proportionate to the risk of harm on their unique platform. It should also account for evolving requirements under other legal regimes, including the potential for age verification to be required at the app store or device level, in order to minimise unnecessary regulatory burden on platform operators.

*32. Do you agree with the scope of the services captured by AA1-6?*

On their face, Measures AA1-6 apply a range of uses of highly effective age assurance to either limit access to a service (or parts thereof) or protect children from encountering harmful content. The applicability of these measures is conditioned on whether a service has as its primary purpose the distribution of content harmful to children, does not prohibit such content, and/or is at higher risk of such content being shared on its service. Conversely, then, these measures imply that a platform which does not meet these conditions would not be required to implement highly effective age assurance.

Like other platforms, Pinterest implements teen safety measures that are not among the measures named in AA1-6, but do require knowledge of a user's age, for example, enforcing its minimum age requirements and defaulting teen users to private profiles. We are concerned, then, that several statements in Volume 5 of the Consultation are broader than Measures AA1-6, and seem to suggest that in practice, *all* platforms will

be expected to utilise highly effective age assurance in order to ascertain the age of their users and provide an age-appropriate experience. For example:

- “We are proposing broader use of age assurance so that services know which of their users are children, so they have a safe experience online – see Section 15. We are proposing ‘age assurance’ to be used by services that pose risks to children. Where we recommend services use age assurance, we propose that they use what we refer to as ‘highly effective age assurance’.” (Vol. 5, p. 10.)
- “Ultimately, our proposals are designed to protect children from encountering harmful content, and to strengthen the effectiveness of other measures we set out in Volume 5 (which might rely on knowing the age of a user).” (Vol. 5, p. 10.)
- “The overarching aim of age assurance measures for services under the children’s safety duties is to help ensure children are protected from harm and receive age-appropriate experiences.” (Vol. 5, p. 34.)

Accordingly, for the sake of providing legal certainty to platforms, it would be helpful for Ofcom to clarify that highly-effective age assurance will only be required for platforms that meet the criteria laid out in Measures AA1-6.

*33. Do you have any information or evidence on different ways that services could use highly effective age assurance to meet the outcome that children are prevented from encountering identified PPC, or protected from encountering identified PC under Measures AA3 and AA4, respectively?*

Regarding this question, we would reiterate Ofcom’s own observation that “age assurance is not a silver bullet and will not be the only effective solution to protect children in all scenarios.”<sup>8</sup> Even if a platform has complete certainty regarding users’ ages, that information does not necessarily lead to ensuring children are prevented from encountering Primary Priority or Priority Content, as required by Measures AA3 and AA4, respectively. Achieving this end is essentially the sum total of platforms’ content moderation efforts, including defining, identifying, and actioning the content at issue. Despite best efforts, this is an imperfect process and one that is not always instantaneous. As such, platforms cannot completely “ensure” that users known to be children will never view harmful content, and this consideration should be taken into account as Ofcom assesses whether Measures AA3 or AA4 have been met.

*34. Do you have any comments on our assessment of the implications of the proposed Measures AA1-6 on children, adults or services?*

---

<sup>8</sup> Summary of Consultation, p. 14.

The enactment of highly effective age assurance measures may add friction to platforms' account registration process, as some users (including both adults and children) may be unwilling or unable to verify their age using the age assurance methodology implemented. This is a burden that would disproportionately fall on newer market entrants that are competing to win user share from established market players.

It is also foreseeable that platforms may make their content safety policies more restrictive in order to avoid application of Measures AA1-6, which would have the impact of limiting adults' ability to view otherwise legal content.

### Content moderation U2U

*36. Do you agree with our proposals? Please provide the underlying arguments and evidence that support your views.*

Echoing our comments in response to the consultation on illegal harms, Pinterest has the following concerns regarding the measures requiring a content moderation system that "allows for swift action against content harmful to children" (Measure CM1), as well as the setting of performance targets (Measure CM3),<sup>9</sup> and encourages Ofcom to pursue a holistic approach which balances the imperatives of promoting teen safety while protecting free expression online.

- CM1. Having a content moderation function that allows for swift action against content harmful to children

We agree that harmful content should be actioned expeditiously and make good faith efforts to do so. While turnaround time is a valuable metric, we do not believe it should be treated as a determinative factor in whether platforms' content moderation systems are effective.

Since the ultimate goal of content moderation is to prevent users' exposure to harmful content, we place more emphasis on how many users saw a Pin before it was actioned. This emphasis on reach is aligned with the functionality of Pinterest's platform, in which content is displayed to users based on their personal interests, rather than prioritising content that was posted recently or that has "viral" status.

In our global Transparency Report, we report data on the reach of Pins deactivated for violating each of our policy categories, calculated as the number of unique users that

---

<sup>9</sup> See Pinterest's response to Illegal Harms Consultation, at Question 16, Measures 4A (swift action) and 4C (performance targets).

saw each deactivated Pin during the reporting period for at least 1 second before it was deactivated. In most cases, this data indicates that very few users view content before it is deactivated.

As speed is not the sole measure of whether a content moderation system is effective, we believe this measure should be applied flexibly, with consideration given to whether platforms are creating a safe environment for users.

- CM3. Performance targets

Similarly to the requirement to take “swift action,” we do not object overall to the setting of performance targets, but would caution against their use as a determinative measure of effective content moderation.

As Ofcom itself alludes to, performance targets for making decisions quickly as well as accurately can be in tension with each other, particularly where a small number of edge cases can skew average turnaround times because additional analysis is required to make an accurate decision. Strictly measuring platforms against performance targets can also create perverse incentives for platforms to over-enforce their policies in an effort to act “quickly” or to reject valid appeals in order to show higher “accuracy” rates, instead of engaging in careful review of nuanced content.

In enforcing this measure, we encourage Ofcom to allow companies to set performance targets that make sense in the context of their own platforms and content moderation systems, and that performance against those targets should be measured with an eye towards the overall success of content moderation efforts, rather than a reductive consideration of whether targets were met.

#### User reporting and complaints

*43. Do you agree with the proposed user reporting measures to be included in the draft Children’s Safety Codes?*

*a) Please confirm which proposed measure your views relate to and explain your views and provide any arguments and supporting evidence.*

*b) If you responded to our Illegal Harms Consultation and this is relevant to your response here, please signpost to the relevant parts of your prior response.*

While the proposals in this section are largely consistent with Pinterest’s existing practices, we believe that certain of the proposals are overly prescriptive. Instead, platforms should be granted the flexibility to provide transparency into their content

moderation systems in a manner that is aligned with their feature set and existing processes. Ultimately, it would be more effective for these measures to be assessed based on the level of transparency they are offering to users, as opposed to strict compliance with the Code as drafted.

In Measure UR3, Ofcom proposes that when users submit a complaint about content, platforms provide an acknowledgement that includes an indicative timeframe for resolution of the complaint, as well as information on what actions the provider may take in response to the complaint and whether the complainant should expect to hear the outcome of their complaint.

We are concerned that the requirement to send an initial acknowledgement of a complaint may be redundant with transparency measures that platforms already have in place, thereby placing additional regulatory burden on companies without an associated benefit to users. For example, Pinterest users have access to its Reports and Violations Centre (RVC), a central, in-product hub where users can view the status of content that they've reported and content restrictions associated with their account. They can also use the RVC to appeal decisions and track their progress.<sup>10</sup> Additionally, if there is an update to the status of a user report, such as a content deactivation or the outcome of an appeal, they will receive a notification via email prompting them to visit the RVC.

The requirement to include in an acknowledgement the actions the provider may take in response to the complaint may also be redundant with resources already provided by platforms. For example, Pinterest publishes information on content moderation measures in a dedicated Enforcement page on its policy site<sup>11</sup> as well as in its Transparency Report.<sup>12</sup> Including this level of detail in communications to users may make these communications less effective by adding to their length and thus discouraging users from reading them, meaning they would not access the essential information about content moderation decisions being made by the platform.

Finally, as stated in Pinterest's response to the Illegal Harms Consultation,<sup>13</sup> we do not believe that providing indicative timelines for resolution of complaints would meaningfully promote transparency. Depending on the complexity of the report at issue, there may be a range of turnaround times that are within industry standards. A given report may be processed later than this average range, or much sooner. We believe that principles of transparency would be better served by allowing users to track the status

---

<sup>10</sup> More information on the RVC can be found in our Help Centre: <https://help.pinterest.com/article/review-your-reports-and-violations>

<sup>11</sup> Accessible at: <https://policy.pinterest.com/enforcement>

<sup>12</sup> Accessible at: <https://policy.pinterest.com/transparency-report>

<sup>13</sup> See Question 15, at Measure 5C.



of their reports and receive notifications as that status is updated. Taking such an approach would also be an opportunity for Ofcom to ease platforms' compliance burden by aligning with measures required by international regulations, namely the DSA, which does not require complainants to be provided with a timeframe for deciding their complaint, but does contain requirements concerning notifications to users on the status and outcome of complaints.<sup>14</sup>

*44. Do you agree with our proposals to apply each of Measures UR2 (e) and UR3 (b) to all services likely to be accessed by children for all types of complaints?*

*a) Please confirm which proposed measure your views relate to and explain your views and provide any arguments and supporting evidence.*

*b) If you responded to our Illegal Harms Consultation and this is relevant to your response here, please signpost to the relevant parts of your prior response.*

Ofcom notes that as part of its future work, it may consider whether to recommend dedicated reporting channels for children. As we noted in our response to the Illegal Harms Consultation,<sup>15</sup> any additional reporting channel imposes costs on platforms for development, maintenance, and staffing, so the value of such a channel should be assessed against that burden. In this case, we believe it would be far more efficient for platforms to ensure that their reporting mechanisms are accessible and easy to use for everyone on their service, including teens, rather than establishing redundant reporting channels for different cohorts of users.

#### Terms of service and publicly available statements

*46. Do you agree with the proposed Terms of Service / Publicly Available Statements measures to be included in the Children's Safety Codes?*

*a) Please confirm which proposed measures your views relate to and provide any arguments and supporting evidence.*

*b) If you responded to our illegal harms consultation and this is relevant to your response here, please signpost to the relevant parts of your prior response.*

Pinterest is aligned with Ofcom's objectives of ensuring that terms of service (ToS) and other publicly available statements are understandable to younger users, and that resources tailored to a younger audience are made available.

We work hard to set the rules applicable to Pinterest's service in a simple, clear language that is accessible to all users, including teens. We also carefully draft our

---

<sup>14</sup> See DSA Arts. 16, 17, 20.

<sup>15</sup> See Question 15, at Measure 5I regarding trusted flaggers.

terms at a level of detail that sets clear expectations for people using our service, while not being so long as to discourage users from reading them. After each section of the ToS, we also provide brief summaries to further clarify the terms provided. For users who want to learn more, we believe it is a best practice to provide links to more detailed policies in specific areas, which are also written in an approachable, succinct style. This approach strikes a balance between providing users with resources and information, while preserving the readability of the ToS.

In addition to making our policies easily accessible, clearly written, concise, and informative, we provide multiple resources which detail our teen safety measures. These include Pinterest’s teen safety hub,<sup>16</sup> as well as Help Centre articles on “Teen safety options”<sup>17</sup> and “Resources for parents and caregivers of teens.”<sup>18</sup>

While we are largely in agreement with Ofcom’s proposals in this section, we are concerned by the proposal that platforms provide summaries of their risk assessments in their ToS (Measure TS3). By potentially adding significant volume to platforms’ ToS, this proposal runs counter to the objective of making policies more comprehensible, particularly since the proposed content is extraneous to the rules governing use of the service.

To ensure terms of service remain clear and readable, we believe it would be preferable for platforms to be permitted to either provide links to their risk assessments, or at least be granted the flexibility to summarise their risk assessments only briefly. In any event, platforms should also not be required to publish any information that might be used by bad actors to circumvent user safety measures.

## **Conclusion**

Pinterest is grateful for the opportunity to contribute to the important work being done by Ofcom in developing the Guidance and Code. We welcome further engagement as part of this process, and if it would be of assistance, we would be happy to elaborate on or discuss any of the points raised in this response or any other matters of interest to Ofcom.

\* \* \*

---

<sup>16</sup> <https://business.pinterest.com/teen-safety-on-pinterest/>

<sup>17</sup> <https://help.pinterest.com/article/teen-safety-options/>

<sup>18</sup> <https://help.pinterest.com/article/resources-for-parents-and-caregivers-of-teens/>