



Dame Melanie Dawes
Chief Executive
Ofcom
Riverside House
London SE1 9HA

27th January 2025

Dear Melanie,

I'm writing in respect of Meta's recent decision to change its Community Guidelines and substantially scale back its proactive moderation of harmful content.

Meta has told us that they are 'not currently able to talk further about the recent company announcement' but based on what is in the public domain we have deep concerns about the impact of this bonfire of safety measures. It appears Meta is returning to the cavalier approach to user safety that has caused untold harm to families across the UK and that directly contributed to Molly's death just over seven years ago.

Put simply, Meta's announcement marks a major strategic recalibration in its approach to protecting users, and this deeply retrograde step is a fundamental retrenchment from the objectives that underpin and are indeed set out in the very first clause of the Online Safety Act.

Mark Zuckerberg's decisions raise profound questions about how soon-to-be regulated companies take decisions about changes to their products, and about how the Online Safety Act might have applied if its relevant provisions were already in force.

Given the fundamental issues that are at stake, MRF considers it essential that Ofcom provides clarity on how it considers the Online Safety Act will apply to these proposed changes. We also believe that it is firmly in the public interest for the regulator to set out how and whether it envisages further changes to its Protection of Children measures will be needed, and to set out the timescales in which necessary strengthening could be made.

We strongly recommend that Ofcom commit to stronger and additional measures to respond to the severity of Meta's proposed rollback of safety measures. In our assessment, these measures should be fast-tracked and built into the first Protection of Children Code before it takes effect in July.

1. The process of arriving at Meta's policy and process changes

Recent media reports suggests that the decision to make such substantial revisions to Meta's approach was taken by Mark Zuckerberg in conjunction with a small number of senior staff.

According to the New York Times, this was a ‘highly unusual’ process that consisted of a six-week sprint exercise, made without any of the usual inputs from internal or external stakeholders.

MRF understands that Meta’s usual process for making changes to its terms and conditions usually takes at least six months. As a result, this announcement was made without any of the usual inputs from internal or external stakeholders.

The process by which Meta’s decisions were taken raise substantial questions about whether this would be compatible with the provisions of the Online Safety Act, if it was already in force. This pronounced deviation from Meta’s pre-existing process it is said to have ‘blindsid[ed] even employees on [Zuckerberg’s] policy and integrity teams’. I would therefore be grateful if you could set out whether consider Ofcom considers this decision-making process to be in line with the provisions in the Act, and if you believe that changes to either Ofcom’s governance measures or the Act itself may be required.

It would be particularly helpful to understand how you would envisage the risk assessment duties would interplay with this expedited process for decision-making, and whether a process which appears to have circumvented the usual process and internal teams could be expected to meet the ‘suitable and sufficient’ standard set out in the Act.

I would also be grateful if you could clarify your understanding of whether you believe the Named Persons measures set out in the relevant Codes would likely be met by similar deviations from Meta’s established process, if these were repeated.

2. The impact of Meta’s policy and process changes on the mental health and well-being of young people, including those experiencing suicidality and self-harm ideation

While Meta has clarified through its press office that prohibited forms of suicide and self-harm content will continue to be treated as a high priority violation and subject to proactive moderation, Molly Rose Foundation nonetheless has grave concerns about the likely impact of Meta’s changes on the mental health, safety and wellbeing of young people.

As Ofcom has itself recognised in the Register of Risks that underpins its draft Protection of Children measures, much of the potential threat to the safety and well-being of young people (including the risk profile relating to suicide and self-harm) comes not only from material that explicitly breaches Meta’s terms of service, but from lower-level forms of content that may have a substantial deleterious effect on children’s mental health and well-being, particularly if algorithmically recommended and consumed in large amounts.

Meta has declined to clarify whether it intends to continue to proactively scan for categories of content that may contribute towards cumulative harm, including but not limited to material that contains non-violative suicide and self-harm content, intense themes of depression, eating disorder content and/or that features dangerous challenges.

In light of the considerable risks that these categories of content can cause when viewed or recommended in large amounts, Ofcom has already proposed to designate much of this material as Non-Designated Content.

It would be helpful if Ofcom could set out its understanding whether Meta intends to continue to proactively scan each of these categories of content, including to moderate and apply the provisions of its Teen Accounts feature; and whether in order to be compliant with the provisions of the Act and meet its expectations, the regulator believes it should.

In light of Meta's announcement, I would also be grateful if you could set out whether Ofcom intends to review and or tighten the relevant recommended measures set out in the draft Protection of Children Code in other ways (and the potential timescales for doing so).

As presently drafted, in most instances relevant regulated firms will need to prevent relevant forms of Non-Designated Content (NDC) from being algorithmically recommended to child users, where relevant information is available to suggest that it is Primary Priority or NDC. However, in practice, these important safety-by-design provisions will effectively be operationalised by content moderation arrangements. Prior to Meta's announcement, MRF had already expressed substantial concerns about the likely effectiveness and robustness of Ofcom's proposed approach.

Meta's regulatory disclosures demonstrate that Instagram and Facebook currently account for less than 2 per cent of suicide and self-harm content moderated by industry as a whole. It is inherently foreseeable that these volumes will fall further (in both absolute and relative terms) if the platform is allowed to shift to a primary reliance on user reports.

3. The impact of Meta's hate speech policy changes on the mental health and wellbeing of young people, and their compatibility with the regulatory scheme

As part of Meta's announcement, Mark Zuckerberg set out significant changes to its hate and prohibited speech policies. This amounts to a substantial rollback of the policies previously in place on Instagram and Facebook, and it is reasonable to expect will result in a significant coarsening of the discourse on Meta's platforms.

MRF is particularly concerned about the impact these changes may have on many younger users, including LGBTQ+ teenagers and young adults, neurodiverse young people, and those from diverse backgrounds. Each of these groups already face heightened exposure to online harms, and each experience higher than average suicide rates.

In this context, it is deeply concerning that Meta will now allow a host of offensive and likely harmful content on their platforms – the suggestion that gay people are mentally ill, and women are 'property,' to name just two.

It seems eminently likely, if indeed not certain, these changes will lead to an increase in bullying among groups who identify with or hold a protected characteristic; that this will negatively impact the right to free expression and association of these groups; and more broadly that these changes may negatively affect the mental health and wellbeing of young users, including young adults.

Given all of this, I would be hugely grateful if you could urgent clarify:

- whether Ofcom envisages these changes if applied to under 18s are consistent with its Protection of Children measures;

- if these are not deemed consistent with either the provisions of the Act or the intention behind Ofcom's draft Codes, what further measures (if any) will be necessary to protect young people's mental health and wellbeing, and to preserve their psychological integrity;
- whether Ofcom will explicitly consider the adverse implications of these changes on the rights to free expression and association of affected groups, and how it will weight these considerations when deciding on its approach.

4. Mark Zuckerberg's comments on 'masculine energy' and the implications for regulatory compliance and conduct

In comments made after the changes to Meta's policies and processes were made, Mark Zuckerberg told the podcaster Joe Rogan that tech companies had become 'culturally neutered'. He stated that he wanted to unleash 'masculine energy' at Meta and added that planned create a corporate culture that 'celebrates the aggression a bit more.'

These comments are clearly deeply concerning. These appear to run entirely contrary to many of the stated objectives in the Act: including but not limited to offering increased protections to women and girls, reducing the risks of offline violence and societal harm, and they inherently contradict the principles and measures set out in the draft Protection of Children Code.

In respect of regulatory compliance, recent weeks have demonstrated the incredibly centralised nature of Meta's strategic and operational decision-making. On this basis, Mark Zuckerberg's comments must surely be taken exceptionally seriously; and when viewed through a regulatory lens, must raise substantial question marks about Meta's commitment to regulatory compliance and embedding safety-by-design.

I would therefore be grateful if you could set out what if any measures you believe are available to you under the Act to address the palpable conduct and behavioural implications raised by Mark Zuckerberg's comments and stated strategic approach.

I would also appreciate you setting out what if any steps you plan to take to address the potential chilling effect of this on the company's approach to moderation, compliance and user safety.

I very much look forward to your response. Give the profound importance of the issues raised, we will be sharing this letter publicly.

Yours Sincerely,

Andy Burrows

Chief Executive