Submission to DCMS Select Committee Immersive and Addictive technologies Inquiry

January 2019

Summary

1. This submission on behalf of Carnegie UK Trust sets out how a duty of care for social media to reduce harms online, developed by Professor Lorna Woods and William Perrin, might be applied to some of the considerations set out in the Terms of Reference for this Select Committee inquiry.

2. Comments on the immersive media industry, the future of eSports, gamification and VR/AR, and the links between gaming and gambling are outwith our remit. Our focus here is on how the "duty of care" approach can provide a framework that applies as much to harms associated with gaming and immersive technology as it can to wider social media harms.

3. For example, we set out below how a "duty of care" would require a "safety by design" approach, through which all online services would be required to be able to demonstrate that they have developed their services with user safety in mind, particularly when those services are used by the most vulnerable, including children.

4. Online behaviours, and the potential harms associated with them, are evolving fast. The uses of immersive technologies are no exception. This presents a challenge to the development of responsive interventions, whether broad or narrow in scope, from regulators and other bodies, while leaving the most vulnerable users of social media and other platforms at ongoing risk of harm.

5. Our approach therefore builds on the well-established "precautionary principle" which has been used by the UK Government since the 1990s to address the threats caused by innovative and novel technologies, when sufficient evidence has not yet been collected to inform standard policymaking considerations.

6. We have kept this submission brief and provided links throughout the document to fuller material for reference. We would be happy to provide more material to the Committee as required.

A statutory duty of care and the precautionary principle

7. Social media and gaming platforms are forms of public spaces. People go to such platforms for all sorts of activities and, while using them, should be protected from reasonably foreseeable harm as they would expect in any public place, such as an office, bar or theme park. While some places are

subject to specific regimes (e.g. pubs), other rules apply more generally, for example the Occupiers Liability Act 1957 and the Health and Safety at Work Act 1974 each of which impose a statutory duty of care. The statutory duty of care set out in the Health and Safety Act contains a number of elements. Section 2 specifies, for example, that:

> it shall be the duty of every employer to ensure, so far as is reasonably practicable, the health, safety and welfare at work of all his employees.

The duty of care is extended to those who are not employees in section 3:

> it shall be the duty of every employer to conduct his undertaking in such a way as to ensure, so far as is reasonably practicable, that persons not in his employment who may be affected thereby are not thereby exposed to risks to their health or safety.

8. This concept is straightforward in principle and well-established; it is also risk-based and future-proofed. A person (including companies) under a duty of care must take care in relation to a particular activity as it affects particular people or things. If that person does not take care, and someone comes to a harm identified in the relevant regime as a result, there are legal consequences, primarily through a regulatory scheme but also with the option of personal legal redress.

9. We set out in our proposals[1] how a statutory duty of care, imposed upon social media companies in respect of their users and enforced by a regulator, would reduce reasonably foreseeable harm caused by social media services. We believe the Committee should consider its applicability in relation to emerging technologies, such as AR and VR, as well as gaming platforms.[2] The boundary between some of these technologies, gaming platforms in particular, and social media is not clear-cut. Adopting a similar approach would support regulatory coherence across the sector. The focus in a statutory duty of care on specifying the outcome, rather than detailing how it happens, lends such duties to rapidly changing and diverse environments such as digital services. It allows those providing the services to address issues as they arise on that provider's service and allows flexibility in response to changing circumstances.

10. The Committee's Terms of Reference ask for views on digital addiction and gaming addiction: "what is the scale of the problem and what support do those with digital or gaming addiction need?". These are good questions: evidence-based policymaking requires that policy decisions should be informed by rigorously established objective evidence. Typically, action on an issue is only taken after consultation and the collection of such scientific or large-scale evidence.

11. But, in innovative areas, there is often no long-term scientific research; or such evidence arrives too late to provide an effective measure against harms. Rapidly-propagating services, such as gaming platforms – which often combine live-streaming with user-interaction akin to that on a social media

---

1 https://www.carnegieuktrust.org.uk/project/harm-reduction-in-social-media/

2 We have been struck by the fact it is a livestreaming gaming platform – Twitch – which has gone further than many social media platforms in baking in user protection from abuse into its Terms of Service and addressing the multi-platform aspect of such abuse, where victims are targeted with co-ordinated attacks via different services. In Twitch's case, users or viewers of the service were using Twitter to abuse or co-ordinate abuse of Twitch users. Twitch changed its TOS to allow it to take action against users on Twitch who are reported to be using other services to abuse Twitch users. Twitch also limits its own users in this regard: 'We prohibit using Twitch to facilitate hateful conduct or harassment, whether the targets are on or off Twitch' https://www.twitch.tv/p/legal/community-guide-lines/ and https://www.bbc.com/news/newsbeat-43003470

platform – are subject to waves of fashion amongst young people and are a particular challenge for long-term objective evidence. Although the body of academic evidence on harms arising from gaming may not be clear-cut, and public perception is mixed [3], there is rising awareness of the "toxicity" of abuse on platforms[4] and the addictive intent of much social media design.[5]

12.   This challenge of inconclusive or incomplete evidence is common in all fast-moving technological areas. For example, while there is a growing body of survey material and qualitative research reports examining the impact of social media on the mental health of children and young people along with research into the potentially harmful cognitive effects, these studies often only demonstrate a correlation, eg between excessive screen time and negative mental health outcomes in children and adults, rather than causation. There is a similar challenge with available research into the impacts of addictive gaming.[6] This leads to a degree of caution in terms of proposed interventions and calls for more research. However, at the other end of the spectrum, the explosion of social media use amongst young people has corresponded with evidence over a similar period of an increase in self-harm and suicidal behaviour.

13.   In the face of such scientific uncertainty, the precautionary principle provides a framework for risk-based harm prevention. After the many public health and science controversies of the 1990s, the UK government's Interdepartmental Liaison Group on Risk Assessment (ILGRA)1 published a fully worked-up version of the precautionary principle for UK decision makers.

> *'The precautionary principle should be applied when, on the basis of the best scientific advice available in the time-frame for decision-making: there is good reason to believe that harmful effects may occur to human, animal or plant health, or to the environment; and the level of scientific uncertainty about the consequences or likelihoods is such that risk cannot be assessed with sufficient confidence to inform decision-making.'[7]*

14.   The ILGRA document advises regulators on how to act when early evidence of harm to the public is apparent, but before unequivocal scientific advice has had time to emerge, with a particular focus on novel harms. The ILGRA's work is still current and hosted by the Health and Safety Executive (HSE) and we commend it to the Committee for consideration as you undertake this consultation.

Safety by design

15.   The duty of care would focus on the systems which such companies would be obliged to put in place as well as their business practices/operational systems, rather than on the content available via the software. This approach parallels the approach in data protection in relation to privacy by design and default, security by design and impact assessments.

---

3   http://www.pewinternet.org/2015/12/15/gaming-and-gamers/

4   https://www.theguardian.com/games/2018/aug/17/tackling-toxicity-abuse-in-online-video-games-overwatch-rainbow-seige

5   https://www.theguardian.com/technology/2018/may/08/social-media-copies-gambling-methods-to-create-psychological-cravings

6   https://www.telegraph.co.uk/news/2018/06/12/addictive-video-games-may-change-childrens-brains-way-drugs/

7   http://www.hse.gov.uk/aboutus/meetings/committees/ilgra/index.htm

16. The Committee also asks: What role does design play in gaming addiction, or the addictive use of social media, and how might that be managed? We would say that design is everything, as Lawrence Lessig[8] argued back in 1999: computer code sets the conditions on which the Internet (and all computers and technologies that run over it) are used. Code is the architecture of cyberspace and affects what people do online: code permits, facilitates and sometime prohibits. It is becoming increasingly apparent that it also nudges us towards certain behaviour and creates the conditions by which individuals can become addicted to online activities or behaviours. Certainly code is not neutral, nor is it pre-ordained. It arises as a result of design choices by software developers. A statutory duty of care would require the developers to think about the likely consequences of their design choices, thus encouraging them to minimise the adverse effects of their technology.

17. When we started out on this work, we focused primarily on the big social media platforms where we assumed that the most harm to users occurred. We have been persuaded that significant risk of harm can arise in smaller services, particularly to vulnerable groups such as children. The same will doubtless be true of emergent technologies – such as AR and VR – especially where there is a community or interactive element. To address such harms while maintaining a risk-managed approach, we now advocate a "safety-by-design" approach applying to all relevant service providers[9]. This is consistent with Lessig's arguments that everything that happens on the Internet is governed by the code that powers it, which puts the responsibility for any harmful outcomes that arise with those who designed the service.

18. Apps and games can increase in popularity so fast that designers should be compelled to bake "safety by design" into their software and business practices from the beginning. Some groups are sufficiently vulnerable (e.g. children) that any business aiming a service at them should take an appropriate level of care no matter what its size or newness to market. Or, to put it another way, even the smallest sandwich shop has to comply with basic food hygiene rules from the day it opens for business.

How the duty of care works

19. Applying the "duty of care" approach to emerging technologies – particularly where applied in an interactive, community setting – as well as the social media sphere has a number of significant benefits:

- It is simple, broadly-based and largely future-proof – expressed in terms of outcome (the prevention of harm) not specifics of process.

- The regulatory approach is essentially preventative, reducing adverse impact on users before it happens, rather than a system aimed at compensation/redress.

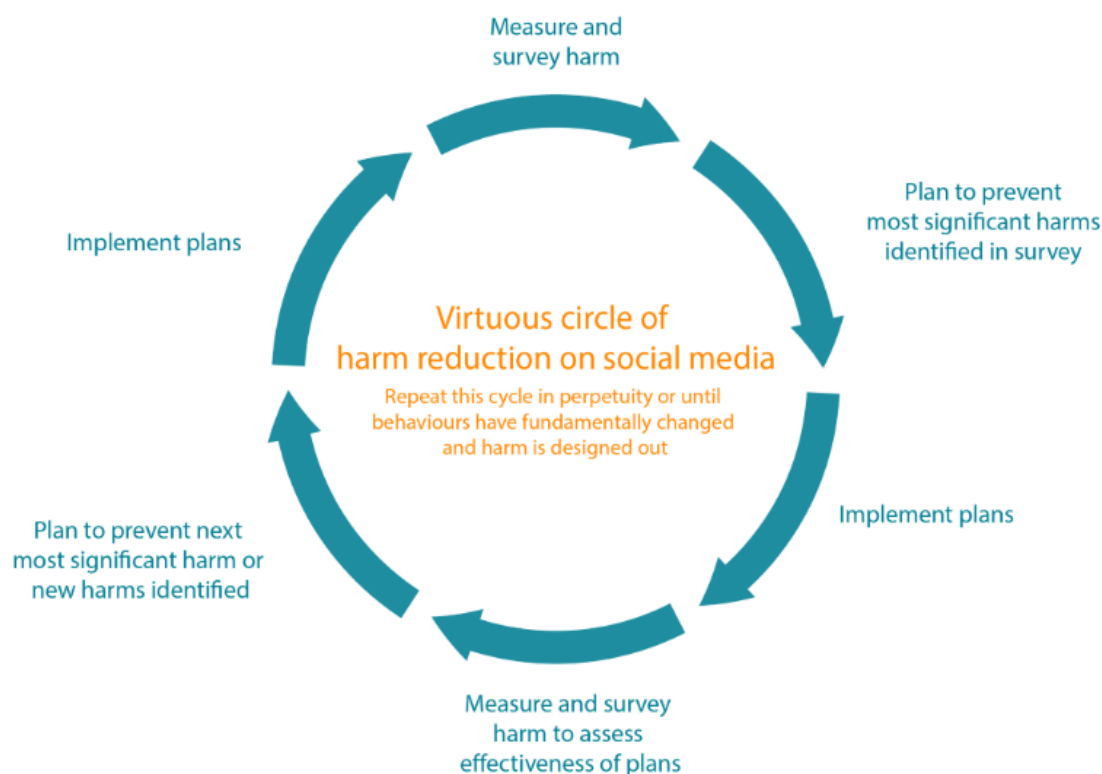- The categories of harm can be specified at a high level, by Parliament, in statute.

---

8   See Lawrence Lessig, "The Law of the Horse: What Cyberlaw Might Teach", (1999),113 Harv. L. Rev. 501 also Code and Other Laws of Cyberspace (1999) and Code: version 2.0 (2006)

9   'For many businesses, all that's required is a basic series of practical tasks that protect people from harm and at the same time protect the future success and growth of your business.' http://www.hse.gov.uk/simple-health-safety/

- A risk-based regulatory approach provides for safe system design (including operational and business choices).

- In micro economic terms returns external costs to the production decision and is efficient if applied in a manner proportionate to risk of harm.

The harm reduction cycle

20. New legislation would set out the duty of care and identify the key harms Parliament wants the regulator to focus on. While not our primary focus, this could include harms associated with gaming and other emerging technologies. In terms of describing the harms, a starting point may be the language found in audio-visual content, such as the revised AudioVisual Media Services Directive[10]. For example, in relation to minors, the impairment of their physical, mental or moral development. This could be adapted in relation to adults, so that the reference could be to the serious impairment of physical or mental well-being.

21. We suggest that the regulator runs a harm reduction cycle,18 as set out in the diagram below, involving civil society as well as companies at each consultative step. The regulator would begin by requiring companies to measure and survey harm, produce plans to address these harms for public consultation and agreement with the regulator then the companies implement the plans. If the cycle does not reduce harms or the companies do not co-operate then sanctions could be deployed.

Measure and survey harm

Plan to prevent most significant harms identified in survey

Implement plans

Virtuous circle of harm reduction on social media
Repeat this cycle in perpetuity or until behaviours have fundamentally changed and harm is designed out

Implement plans

Plan to prevent next most significant harm or new harms identified

Measure and survey harm to assess effectiveness of plans

---

10   We share Baroness Grender's view (in a Lords debate on a social media duty of care) that competent regulators have had little difficulty in working out what harm means: 'If in 2003 there was general acceptance relating to content of programmes for television and radio, protecting the public from offensive and harmful material, why have those definitions changed, or what makes them undeliverable now? Why did we understand what we meant by "harm" in 2003 but appear to ask what it is today?' https://hansard.parliament.uk/Lords/2018-11-12/debates/DF630121-FFEF-49D5-B812-3ABBE43371FA/SocialMediaServices?highlight=social%20media%20services#contribution-CF09F315-56F2-49AE-9F88-ADCAACD3BE21

Oversight and implementation

22. The Committee also asks for views on how the official bodies such as the Office of Artificial Intelligence and the Centre for Data Ethics and Innovation will co-ordinate and share their work. Again, this is not a direct concern of our work but we believe it is worth stating that it takes a long time for new organisations to establish themselves – both in an operational sense as well as in a credible, impactful sense.

23. In the meantime, new harms emerge with increasing regularity on new platforms and services. The government, OFCOM, the ICO and countless lobby groups have described serious harms occurring now apparently at a population scale. We can observe new networks and platforms being created at a rapid rate; many quickly generating new types of harms and threats, particularly to children and young people. While barriers to entry are higher in gaming (particularly using VR/AR) than other forms of social media and technology, capital is plentiful, technology costs are falling still and incentives for self-regulation are non-existent. In considering structural regulatory options, weight should be giving to doing things quickly.

24. That is why we argue that a duty of care be rapidly introduced to deal with the most prevalent harms online and in the immersive and gaming worlds. With Government backing, a short Bill – around 30 clauses – could be introduced and debated, even within a Parliamentary context that is dominated by Brexit uncertainties. Regulatory oversight should be assigned to an existing organisation – for our proposals for a duty of care for harm reduction on social media, we suggest Ofcom, so that the regime can begin quickly to have an impact.

25. We are happy to elaborate further on any of the material above.

William Perrin

Professor Lorna Woods