Online Platforms & Mental Health
A Policy Proposal

Maria Carnovale and Samuel A. Ramirez
Carr Center
Discussion Paper
Online Platforms and Mental Health
A Policy Proposal

Carr Center for Human Rights Policy
Harvard Kennedy School, Harvard University
September 15, 2022

Maria Carnovale
Technology and Human Rights Fellow
Carr Center for Human Rights Policy

Samuel A. Ramirez
Research Associate, Department of Pharmacology & Computational Medicine
University of North Carolina at Chapel Hill

The views expressed in the Carr Center Discussion Paper Series are those of the author(s) and do not necessarily reflect those of the John F. Kennedy School of Government or of Harvard University. Faculty Research Working Papers have not undergone formal review and approval. Such papers are included in this series to elicit feedback and to encourage debate on important public policy challenges. Copyright belongs to the author(s). These papers may be downloaded for personal use only.

1 Technology and Human Rights Fellow, Carr Center for Human Rights Policy, Kennedy School of Government, Harvard University, carnovalemaria@outlook.com

2 Research Associate, Department of Pharmacology and Computational Medicine Program, University of North Carolina at Chapel Hill, s_o_ramirez@outlook.com.
ABSTRACT

In recent years, there has been growing concern regarding the unintended mental health impact of online platforms and whether they might be driving a public health crisis, especially among children and teens. There is emerging evidence that spending too much time on digital platforms—like gaming sites, online pornography sites, and social media—can be associated with negative mental health effects such as depression and social anxiety, at least in some users. Yet most policy action and advocacy in this industry have focused on the issues of privacy and misinformation, relegating the mental health impact of digital technology as a secondary byproduct of the industry. In this paper, we provide an overview of the documented negative mental health effects associated with prolonged use of video games, online pornography, and social media. We outline the measures that have been taken to address the mental health impact of these technologies. Finally, we suggest that induced overuse is at the heart of the problem and we propose an incentive-based policy mechanism to address it.
I. Introduction

In recent years, there has been growing concern regarding the unintended mental health impact of online platforms and whether they might be driving a public health crisis (Christ 2018; Lee Bouygues 2021; Twenge 2017), especially among children and teens. In 2019, before the COVID-19 pandemic, one in three high school students reported persistent feelings of sadness and hopelessness. These numbers increase by 44 percent compared to 2009. Teen suicide and attempted suicide rates have increased by 40 percent in the same decade (Centers for Disease Control and Prevention 2020).

In the 2021 report “Protecting Youth Mental Health,” the US surgeon general called on the role of technology companies, expressing that “when not deployed responsibly and safely, these tools [technology platforms] can pit us against each other, reinforce negative behaviors like bullying and exclusion, and undermine the safe and supportive environments young people need and deserve” (Murthy 2021). The use of digital technologies has improved our lives in important ways. Former UN rapporteur on freedom of expression, Frank La Rue, stated in 2014 that access to the internet is an indispensable tool for children in claiming their freedom of expression as well as other rights, like the rights to education, association, and full participation in social, cultural and political life (Byrne and Kardefelt-Winther 2016).

While providing important opportunities, there has been increasing recognition that digital technologies can also deteriorate the social environment in which children grow and impair the ability of young users to reach their full potential. Sean Parker, the founding president of Facebook, said that the founders of social networking sites are “exploiting vulnerabilities in human psychology” through a validation feedback loop to constantly attract people’s attention. He famously concluded: “God only knows what it’s doing to our children’s brains” (Allen 2017).

In this paper, we focus on the mental health impact of three widespread internet activities: internet gaming, online pornography, and social media. While other internet activities like internet shopping or video streaming might deteriorate the well-being of users, we restrict this paper to the analyses of internet gaming, online pornography, and social media because these three industries are very widespread and present similar characteristics that can be addressed with a unified policy approach.

There are numerous valuable individual and social outcomes from these activities. Video games can decrease the level of stress, relieve aggressiveness, and—in the case of multiplayer gaming—promote teamwork and social interactions. Social media can trigger meaningful social connections, provide a channel for self-expression, and mobilize action for causes that are salient to one’s identity. Online pornography can offer a channel to discover unexplored aspects of one’s sexuality. Nevertheless, there is emerging evidence that spending too much time on these digital platforms can be associated with negative mental health effects such as depression and social anxiety, at least in some users. Internet gaming was officially recognized as an addictive activity by the World Health Organization (WHO) in 2019. These platforms provide Novelty and constant hyperstimulation, emotional and/or sensory, which can rewire brain circuits and increase the users’ dependence on the sense of reward they provide. Young people are more vulnerable because full brain development is not complete until around twenty-five years of age (Arain et al. 2013).

Constant hyperstimulation and novelty, however, are an integral part of digital companies’ business model. Those features drive user engagement, which is a key ingredient in digital companies’ profit streams since their revenues are driven by selling advertisement space. The vast availability of user data allows companies to improve the ways that user attention is harvested and sold.

Yet most policy actions and advocacy have primarily focused on the issues of privacy and misinformation, relegating the mental health impact of digital technology as a secondary by-product of the industry. In this paper, we provide an overview of the documented negative mental health effects associated with prolonged use of video games, online pornography, and social media (Section 2). We then outline the measures that have been taken to address the mental health impact of these technologies (Section 3). Finally, we suggest that overuse is at the heart of the problem, and we propose an incentive-based policy mechanism to address it, discussing its benefits and limitations (Section 4).

Most attempts to regulate the digital industry for its connection to the deteriorating mental health of users have been in the form of a command-and-control policy. China has recently introduced a time limit of three hours a week for minors playing video games (Pruitt-Young 2021). In the US, the proposed bill Social Media Addiction Reduction Technology Act of 2019 suggested banning some of the features of social media that drive prolonged engagement. Both strategies express a concern regarding the overuse of internet gaming and social media through an arguably paternalistic approach that might ultimately limit their effectiveness. Both users and companies, in fact, might adapt to circumvent those constraints.

An incentive-based policy approach might outperform a command-and-control one. Therefore, we propose imposing a tax on technology companies when users spend more time on the platform than a set time limit. This economic incentive might motivate companies to use their behavioral strategies to drive user engagement but discourage overuse. Tax revenues should be reinvested into funding mental and digital health educational programs, sponsoring counseling and other men-
tal health resources, and supporting research into the mental health externalities of digital technologies.

Digital technologies will continue to develop new ways to stimulate users’ senses and capture their attention. In November 2021, Facebook’s founder Mark Zuckerberg announced the rebranding of the company to Meta to reflect the new company’s objective of bringing the metaverse to life (Meta 2021). The metaverse is a combination of augmented and virtual realities that provides for immersive online interactions and will likely usher in a new virtual economy (Ravenscraft 2021). While there will be undeniable benefits, it is important to recognize its mental health risks and proactively regulate these industries to protect those who are most vulnerable.

2. The Mental Health Impact of Internet Gaming, Online Pornography, and Social Media

This section reviews the effects of three types of online activities—video games, online pornography, and social media—on individual users. Of these, only internet gaming addiction and compulsive use of online pornography have been officially recognized as mental health disorders by the WHO in its International Classification of Diseases (ICD-11). While there is suggestive evidence that heavy use of social media is associated with some mental health symptoms, the research is at a relatively early stage. Nevertheless, these industries present similar pathways to attract and retain user attention. For all three types of online activities, scientific inquiry has focused on determining their addictive nature and symptoms, but there is little research on the long-term impact.

2.1 Internet Gaming

On May 25, 2019, the WHO board unanimously voted to include internet gaming addiction in its official list of diseases, the ICD-11 (Kamenetz 2019). The WHO decision followed the 2013 listing of Internet Gaming Disorder in section III of the Diagnostic and Statistical Manual of Mental Disorders—V (DSM-5), a comprehensive list of mental health disorders of the American Psychiatric Association (APA) (Fauth-Bühler and Mann 2017; Jo et al. 2019).

Despite three decades of accumulated research on this topic, this official recognition did not enjoy unified support. The first case studies on video game addiction emerged in the early 1980s, roughly ten years after the introduction of the first commercial video games (Keepers 1990; Klein 1984; Nilles 1982; Ross, Finestone, and Lavin 1982; Soper and Miller 1983). These studies were initial descriptions of symptoms of addiction following continuous video game engagement, such as compulsive...
behavioral involvement, lack of interest in other activities, dissociation with peers and association with other addictive gamers, and withdrawal when attempting to quit. These observations were later supported by subsequent studies (Eichenbaum et al. 2014; Eickhoff et al. 2015; Griffiths 2009; Lehenbauer-Baum et al. 2015; Pontes 2017; Van Rooij et al. 2014; Schmitt and Livingston 2015).

More recently, this research has investigated the neurobiological effects of consistent engagement with video games. Mostly thanks to the development of neuroimaging techniques, several studies have shown that, much like substance use, video games can lead to changes in areas of the brain leading to impaired impulse control, decision-making, behavioral inhibition, emotional regulation, learning and memory, and reward processing (Dong et al. 2012; Dong and Potenza 2014; Fauth-Bühler and Mann 2017; X. Lin, Dong, et al. 2015; X. Lin, Jia, et al. 2015; Liu et al. 2016; Palaus et al. 2017; Pontes, Kuss, and Griffiths 2017; Weinstein, Livny, and Weizman 2017; Weinstein 2017).

Despite this evidence, the decisions of the WHO and the APA were considered controversial because most supporting studies fell short of accepted methodological standards. For instance, studies vary in the core addiction indicators used and the cut-off scores to determine whether a user is considered addicted or not. The empirical evidence is correlational, but it does not establish a causal relationship between video game use and mental health effects.

The lack of a causal link between internet gaming and mental health symptoms is a particularly limiting factor since compulsive behavior and video game overuse might be symptoms of other disorders such as depression or anxiety. Nevertheless, standard causal estimation techniques are difficult to apply to the study of these technologies. For instance, in randomized controlled trials, the study group and the researcher must ignore which subjects are in the control group and which are in the treatment group, but that information is impossible to hide if the treatment is to play or stop playing video games.

Finally, video game addiction is not widespread. A systematic meta-analysis of this literature found that around 3 percent of gamers develop symptoms, even though the authors find high levels of variance attributed to sample characteristics, such as culture and age, and methods of analysis (Stevens et al. 2021). These results are consistent with other studies (Pontes 2018; Wittek et al. 2016) and translate into a prevalence of 0.3–1 percent in the general population. Nevertheless, rates can be much higher among children and adolescents (Gentile et al. 2017; Strittmatter et al. 2015; Ustinavičiene et al. 2016; Wartberg, Kriston, and Thomasius 2017; Yu and Cho 2016).

For those who do develop symptoms, however, a wide range of harmful effects has been associated with excessive gaming and addiction. Those effects span from depression and social anxiety (Brunborg, Mentzoni, and Frøyland 2014; Lehenbauer-Baum et al. 2015) to sleep deprivation (Eichenbaum et al. 2014); lower engagement with school and academic performance (Brunborg, Mentzoni, and Frøyland 2014; Van Rooij et al. 2014; Schmitt and Livingston 2015); less physical exercise (Henchoz et al. 2015); impaired emotional and behavioral functioning (Baer, Saran, and Green 2012; Van Rooij et al. 2014); lower levels of social identity, self-efficacy, and satisfaction with life (Festl, Scharkow, and Quandt 2013); increased levels of stress and aggression (Festl, Scharkow, and Quandt 2013; Scharrer et al. 2018; Snodgrass et al. 2014); and overall poor mental health (Vuksosavljevic-Gvozden, Filipovic, and Opacic 2015; Wittek et al. 2016).

These effects can be destructive to people’s lives, and especially in the case of children and adolescents, they can affect individuals’ long-term development and life trajectory. Nevertheless, unlike the case of online gambling (or other addictive substances like alcohol use and tobacco), the video game industry has mostly escaped government constraints (Van Rooij et al. 2010).

### 2.2 Online Pornography

While the effects of internet gaming on mental health have been widely discussed in mainstream media, more silent to the ears of the general public is the effect of online pornography on users’ mental health and well-being. By now, a substantial body of research has connected the use of online pornography to symptoms of addictions both at a clinical and neurobiological level (Brand et al. 2011; Kühn and Gallinat 2014; Laier et al. 2013; Laier, Pawlikowski, and Brand 2014; Laier, Schulte, and Brand 2013; Steele et al. 2013; Voon et al. 2014), even though the same limitations of the gaming addiction studies apply. In 2018, the WHO included in the ICD-11 (World Health Organization n.d.) a diagnosis for “Compulsive sexual behavior disorder” (listed as 6C72) that includes...
the repetitive use of pornography among its clinical features. Neither the WHO nor the APA has officially recognized addiction to online pornography as a mental illness.

The use of online pornography is extremely widespread. A study conducted in 2015 surveyed 941 young Australians and found that 100 percent of men and 82 percent of women had viewed pornography in the month before the study. Of the respondents, 23 percent of women and 85 percent of men watched pornography weekly or more frequently (Lim et al. 2017). These results can likely be generalized to other countries where sexually explicit media are widely available and where the public has little information on their potential impact on mental health.

According to traffic statistics, in November 2021, the three most common pornography websites combined (Pornhub.com, Xvideos.com, Xnxx.com) were visited more often than Wikipedia and around half as much as either Facebook or YouTube. The traffic of Pornhub.com alone was comparable to that of Twitter, Yahoo, or Reddit (Clement 2022). The popularity of pornographic material is corroborated by the frequency of internet searches. As revealed by Google Trends, within the last year, almost as many people have searched the word “porn” as the word “facebook” or “youtube” worldwide (figure 1).

The magnitude of the mental health disorders associated with porn use is hard to estimate. Nevertheless, online social media are providing suggestive evidence of how many people might be experiencing symptoms based on the number of people seeking support. Online forums and support groups, such as “r/NoFap” on Reddit.com, forum.nofap.org, and forum.rebootnation.org, have sprouted on the internet as a growing number of online pornography users come together to share their experiences and the negative impact of pornography on their mental health. As of March 2022, the largest forum, Reddit’s NoFap (https://www.reddit.com/r/NoFap/) had close to one million users. Most of them appear to be located in North America and the UK.

The users of these forums report physical symptoms such as erectile dysfunction and delayed ejaculation, decreased sensitivity to visual and physical stimulation, and fatigue. There are also psychological symptoms such as low self-esteem, lack of motivation, disinterest in physical intercourse in favor of watching pornographic material, and clouding of consciousness or brain “fog” (NoFap n.d.). These reports are in line with a body of scientific literature linking the use of pornography to sexual dysfunctions (Blair 2017; Bronner and Ben-Zion 2014; Park et al. 2016), poorer relationship satisfaction (Minarcik, Wetterneck, and Short 2016; Muusses, Kerkhof, and Finkenauer 2015; Perry 2017a, 2018; Perry and Schleifer 2017; Wright et al. 2017), and depressive symptoms (Cooper et al. 2011; Nelson, Padilla-Walker, and Carroll 2010; Owens et al. 2012; Perry 2017b; Peter and Valkenburg 2016; Yard et al. 2022; Yoder, Virden, and Amin 2007).

2.3 Social Media

Social media can nurture communities by connecting people with common characteristics or ideas. They can foster a sense of purpose and empowerment by allowing users to voice their

![Daily worldwide searches of the word “porn” (red line), “facebook” (blue line), “youtube” (yellow line), and “pornhub” (green line) between January 6, 2019, and March 26, 2022. Source: Google Trends.](attachment:image.png)
work, beliefs, and experiences. They can connect those in need with mental health resources.

Nevertheless, there is growing evidence showing that social media platforms can have negative effects on users. Studies have shown that social media use is connected to poor sleep (Royal Society for Public Health 2017; Woods and Scott 2016), depression, anxiety, and social isolation (L. Y. Lin et al. 2016; Primack et al. 2017; Woods and Scott 2016). It can expose users to bullying and it can instigate lack of self-esteem (Ozimek and Bierhoff 2019; Woods and Scott 2016). Heavy users seem to be the most unhappy after interacting with social media, according to data collected by Moment, an activity-tracking application (Data Team 2018).

While the connection between social media use and mental health effects might be small in magnitude within the general population (Resnick 2019), some subgroups appear to be more vulnerable. Social media use can be an important mediator for depression among sexual minorities (Escobar-Viera et al. 2018), and negative effects are stronger among adolescents (Keles, McCrae, and Grealish 2019)—especially among young girls (Viner et al. 2019).

Yet, the causal link is still an open question: we still don’t know whether social media weaken users’ emotional stability or if emotionally vulnerable individuals are prone to engage with social media for longer times or in self-harmful ways. Like the cases of internet gaming and online pornography, the barriers to designing studies that apply standard causal methodologies make it hard to establish (or disprove) a causal link beyond the anecdotal evidence provided by guardians or educators who often observe changes in the personalities and moods of children and teenagers following extensive use of social media.

A related question is whether social media use can be addictive. After all, it does share many characteristics with video games, like the opportunity for a repetitive and short-term sense of reward. Since social media companies’ profits are based on advertisement revenues, those platforms are designed to draw users to the platform and make it hard to look away (Wägström 2018). This was very clearly explained by Sean Parker, former president of Facebook, in an interview for the magazine Axios (Pandley 2017): “The thought process that went into building these applications, Facebook being the first of them, was all about: ‘How do we consume as much of your time and conscious attention as possible?’ . . . And that means that we need to sort of give you a little dopamine hit every once in a while, because someone liked or commented on a photo or a post or whatever. . . . It’s a social-validation feedback loop” (Allen 2017).

Systematic reviews of the literature suggest that social media might, in fact, become addictive for some users (D’Arienzo, Boursier, and Griffiths 2019; Kuss and Griffiths 2011, 2017). In a 2017 survey, the Dutch Central Bureau of Statistics found that 10.1 percent of the Dutch population feels addicted to social media. Numbers are much higher at younger ages: short of 29 percent of respondents between eighteen and twenty-four years of age and 16.7 percent of those between twenty-five and thirty-four years old self-defined as addicted to social media. These numbers are consistent with self-reported symptoms of addictions: the group of respondents between eighteen and twenty-four years of age reported negative effects on their sleep (40%) and academic performance (35%) as well as mild feelings of withdrawal such as discomfort (36%) and restlessness (21%) when unable to go online. Numbers were higher compared to the 2015 version of the survey (Central Bureau of Statistics 2018). Using neuroimaging techniques, a study showed that Facebook use can impair impulse control similar to how gambling and substance use do (Turel et al. 2014). Nevertheless, the dispute over whether social media are addictive per se or if social media overuse is a symptom of other disorders is still open.

3. On-the-Ground Responses and Recommendations

There have been many approaches to address the mental health effects of internet gaming, online pornography, and social media. Responses and proposals include advocacy for official recognition by public health institutions of the mental health disorders associated with the use of those internet platforms, the setup of educational programs, calls for companies’ voluntary interventions, and governmental interventions. All of these present benefits and drawbacks. Given the complexity of the economic and social implications of these technological services/products, all of these interventions are important components of the overall social response. In this section, we provide an overview of these four approaches and provide recommendations.

3.1 Inclusion into the ICD and/or DSM

To achieve official recognition of the negative health effects and especially the addictive nature of some of these technologies, many organizations actively advocate for the inclusion of the associated mental health disorders into the ICD or the DSM. By doing so, activists seek to increase the resources dedicated to the issue. Official recognition of mental disorders is also meant to gain support for policy interventions against those indus-
“When not deployed responsibly and safely, these tools [technology platforms] can pit us against each other, reinforce negative behaviors like bullying and exclusion, and undermine the safe and supportive environments young people need and deserve.”

— US Surgeon General,
2021 report “Protecting Youth Mental Health”
tries that thrive despite the negative health effects imposed on heavy users, in much the same way that the recognition of the negative health effects of smoking cigarettes has spurred policy interventions in the tobacco industry. Destigmatizing mental health disorders is another benefit of inclusion in the ICD and/or the DSM. Finally, the act of advocating for an official recognition might sensibilize public opinion, inform people, and point affected users to helpful resources.

This process has led to the inclusion of internet gaming among the list of addictions. Many organizations like Your Brain on Porn have, for a long time, demanded the recognition of watching online pornography as an addictive activity as well. Advocacy concerning the risks of social media use is still at an earlier stage possibly because of the relative novelty of these platforms.

This approach can sometimes backfire. For instance, organizations advocating for the recognition of addiction to online pornography are often confused with religious fanaticism or the imposition of a heteronormative view of individual sexuality. Internet gaming addiction, and now social media overuse, is often portrayed as a form of public hysteria and resistance to change, comparable to when, in the 1800s, modern novels were considered to overexcite young women’s minds and have pernicious effects on their mental health (North 2014; Oxford Royale Academy n.d.).

This pushback might be driven by the sensationalistic media outreach that accompanies advocacy. In part, however, this is simply an effect of a dispersed scientific debate. The case of video games has much to teach: it took almost 30 years before the evidence collected led to official recognition by the WHO. Even after its official recognition, there has been limited social or policy action addressing the gaming industry.

3.2 Educational and Informational Interventions

There is a wealth of digital citizenship programs designed for school-age kids. The content teaches users how to engage on internet platforms appropriately and responsibly and educates children on how to use the internet in a way that preserves their own health and that of others. These programs provide information to children, parents, and educators. Some examples of organizations in this space are Common Sense Media and the LEGO Group’s Raise Digitally Smart Families program. Brain POP provides educational content for educators and schools. Many universities are also adding digital citizenship programs to their curricula.

While this is a good start, these programs often prime users to accept internet platforms for what they are and adapt users’ behavior to minimize harm rather than critically rethinking the way internet platforms are structured. They also pose an additional burden on families and educators who need to devote time and resources to researching and implementing these programs within their communities. This is an investment that might increase inequality in individual mental health outcomes based on income differentials.

Van Rooij and coauthors also suggest that internet gaming sites should include warnings that inform the users of the potential health risks and provide referral services to institutions that might help if the user were to experience any symptoms (Van Rooij et al. 2010). In doing so, they suggest an intervention similar to the cautionary labels on tobacco products or online gambling sites.

3.3 Companies’ Self-Regulation

With technology changing fast, advocates of companies’ self-regulation support it for its flexibility and quick reaction to events, social phenomena, and public opinion. So far, in these industries, companies’ self-regulation has been mostly limited to ethical design for video games (design that limits any addictive feature) and content moderation for social media.

Content moderation is the process of screening for and eliminating or flagging users’ posts on social media when their content does not satisfy some pre-set guidelines. Those guidelines can be set by government regulation. This is the case when the content is removed because it was violating some intellectual property or was the subject of a request of the third party based on the right to be forgotten (in the European Union). But it can also be a voluntary company practice. For instance, after the Cambridge Analytica scandal in which Facebook was allegedly leveraged to influence the Brexit vote in the UK and the 2016 presidential election in the US, the potential of social media to be used as a tool for social manipulation through targeted and often false content gained prominence. Many social media platforms like Facebook or Twitter have since introduced mechanisms to flag or delete user content that is false or excessively divisive or that promotes hate speech, even though these systems have varied in their longevity (Cellar-Jones 2017; Clayton 2021).
Content moderation has also been suggested as a tool to curb cyberbullying by taking down content that targets a single individual with threats or derogatory content. Cyberbullying—a form of violent verbal attack and life threat that is relatively common, especially among children, adolescents, and young adults (Hamm et al. 2015; MacDonald and Roberts-Pittman 2010)—has been connected to symptoms of depression, suicidality, and substance use (Hamm et al. 2015; Selkie, Fales, and Moreno 2016). It can lead to severe impacts. In 2017, thirteen-year-old Rosalie Avila hung herself after two years of bullying that debilitated her sense of self-worth (Hansen 2017). The same year, Ashawnty Davis died by suicide after a video of herself went viral on the social network Musical.ly. She was only ten years old (Segarra 2017). Nevertheless, the response by social media companies has been scant and taken on a case-by-case basis following a user request. A more systematic approach might limit instances of verbal violence against individuals and their tragic consequences.

Still, over-relying on companies’ voluntary interventions might ultimately be inconclusive when addressing the mental health issues stemming from prolonged use of internet platforms. Driving continuous user engagement, in fact, is not a feature of internet platforms, but their business model. Their revenues come from providing advertisement space to companies that pay for the service only if a large set of viewers are present on the platform to see those ads. Aggressive, offensive, and divisive content drives more user engagement than balanced and respectful posts. As a result, that content is often more visible not only because more users react to it but also because social media algorithms highlight it for the audience most sensitive to it (Lanier 2018).

Social media companies collect fine-grained behavioral data about their users. They use those data to find the emotional weaknesses that drive engagement with the platform. The larger the amount of time spent on the platform, the larger the amount of data on each individual, and the more precisely the algorithm can predict what content will “push the buttons” of its users and feed them content that will cater to that response. In other words, social media algorithms emotionally destabilize users by presenting content that constantly drives engagement. Aggressive content often has that effect and is therefore rewarded (Lanier 2018).

Relying on voluntary interventions on the part of social media companies to protect users against potentially developing mental health issues might be unreasonable because it violates the way those companies generate profits. This is very different from self-regulation on a secondary feature of a product, like decreasing the environmental impact or labor abuses in the delivery of consumer products like food or apparel, which is where voluntary standards and self-regulation have been commonly applied. Asking companies that manage internet platforms to self-regulate is no different than asking tobacco companies to stop poisoning their customers without any regulation to trigger that change in behavior.

### 3.4 Governmental Regulation

While the policy debate has mostly focused on privacy (and misinformation for social media), some attempts to regulate internet gaming, online pornography, and social media use in recognition of their mental health impact, especially on children and teens, have already been taking place. In the UK, plans to require pornographic websites to verify the age of the users have been met with substantial criticism based on limiting individual privacy on the internet (Milmo and Waterson 2022). China has recently introduced a time limit of three hours a week for minors playing video games (Pruitt-Young 2021).

In June 2019, Senator Josh Hawley introduced the Social Media Addiction Reduction Technology Act (or SMART Act) intending to curb features of social media platforms that encourage prolonged engagement. The bill, which did not pass, proposed to ban features such as infinite scrolling, auto-play, and awards for social media engagement from online applications and websites. It also proposed to require companies to include in-application tools for users to monitor and limit the amount of time spent on the platform, with a 30-minute limit across devices as the default option.

The SMART Act proposal generated much debate. Supporters like Josh Golin (Executive Director of Campaign for a Commercial-Free Childhood) commended the effort on the basis that “social media companies deploy a host of tactics designed to manipulate users in ways that undermine their wellbeing” (Senator Josh Hawley 2019). Its critics, however, pointed to the lack of scientific consensus on the addictive nature of social media platforms. In a statement, Michael Beckerman (former President and CEO of the Internet Association) said: “Policy proposals must be evidence-based, which is why the internet industry supports the CAMRA [Children and Media Research Advancement] Act. Funding for independent scientific research is critical to better understand
the impact of screen time and identify more ways to enhance people’s well-being on and offline” (Rogers 2019).

While many questions remain open, there is a risk in the well-known formula “more research is needed.” It can lead, as in the case of video games, to thirty years of ultimately inconclusive work due to the previously mentioned limitations in study methodologies. Sound scientific research is an important basis for policy action, but scientific inquiry that supports evidence-based policies is by nature narrow, incremental, and most importantly slow. In these industries, innovation advances rapidly with new ways to capture and retain people’s attention. The scientific research that assesses technology’s effect on users’ mental health will constantly be playing catch up with its innovations and therefore produce consistently outdated knowledge. Ultimately, a more productive strategy might be to adopt a precautionary approach and shift the burden of the proof on the targeted companies.

While this paper outlined the negative mental health effects of engaging with internet gaming sites, social media, and online pornography, it is important to remember that these platforms also produce many positive social outcomes. Video games can promote teamwork and social interactions. Social media can support meaningful social connections. Online pornography can be a form of sexual expression. What internet users who experience problems appear to have in common is their prolonged engagement on these platforms.

Nevertheless, many features of social networking sites are designed to increase the time that users spend on those sites. Video suggestions and automatic video queues on YouTube, infinite scrolling on social media, and notifications to bring you back to the site or mobile applications are some examples of these strategies.

From a policy perspective, therefore, the question is how to preserve the positive outcomes of user engagement with internet platforms while curbing the negative impacts. The proposal of the SMART Act of 2019 and China’s time limits for video games are attempts at achieving that goal. Nevertheless, the regulated companies might eventually find ways to achieve the same objectives around the limitations imposed.

4. Proposed Intervention

An incentive-based approach might be better suited because it provides monetary incentives for companies to adapt to the target outcome, leaving them with the freedom to choose the best way to achieve that outcome. We recommend taxing the technology companies that own the online platforms for every user who spends an excessive amount of time on the platform. The tax should be imposed on companies, not users. We propose the following functional form to compute the tax:

\[
T = \sum_{i=1}^{N} t(x_i) \text{ where } t(x_i) = \begin{cases} 
0, & x_i < \bar{x} \\
f(x_i - \bar{x}), & x_i \geq \bar{x}
\end{cases}
\]

The tax should be calculated by setting a time limit below which engagement with the platform is not considered a harmful behavior (the time limit is \(\bar{x}\) in the formula above). For each user \((i)\) whose time spent on the platform or online application \((x_i)\) exceeds the set time limit \((x_i \geq \bar{x})\), the platform should pay a tax computed based on the spread between the observed user behavior and the threshold \((f(x_i - \bar{x}))\). When users spend less time than the threshold on the platform, the tax will be equal to 0. The company’s tax liability should be the sum of the individually calculated taxes across all users who reside within the regulating jurisdiction (with \(N\) being the total number of users in the regulating jurisdiction).

This policy mechanism provides an incentive for companies to use their behavioral strategies and the data they collect to promote engagement but discourage compulsive use. Companies’ responses might take place within or outside of the online platform. For instance, it might drive companies to stop notifications that entice users to rejoin the platform once they are close to the time limit. An alternative response might be to provide funding for offline programs and activities that educate users about the mental health risks of prolonged engagement with the online platform.
In the Pigouvian tradition, the functional form to calculate the tax \( f(x_i) \) should reflect the social cost of the behavior or externality that the regulation aims to discourage. In this context, however, setting a traditional Pigouvian tax implies calculating the monetary value of the potential mental health impact that an individual might suffer through excessive use of internet platforms and the effect of these platforms on society as a whole. While there are valuation methods to compute it, the process would likely be long and the result controversial. Therefore, the suggested approach is to impose a tax that is proportional to the revenues that social media companies get from an extra time unit of average user engagement. By offsetting the profits made through advertising, the regulated companies should have a strong incentive to use their resources toward having their users comply with the time limit.

Crucially, this regulation should contain language that forbids companies to pass the tax on to the users. For instance, companies should not be allowed to charge a fee to users exceeding the time limit or offer a for-pay premium membership that allows for extended use. Similarly, companies should not be allowed to set barriers for over-users, like shutting down the site or introducing bugs in its functioning. These responses, in fact, would likely incentivize users to create multiple accounts and switch between them to evade the time limit.

We also recommend revising the tax at regular intervals, for instance every five years. This is very common with environmental regulation to curb the negative environmental impact of economic activities because it provides policymakers with the flexibility to address fast-paced industries. It also allows for an assessment of the performance of the tax instrument at regular intervals and making any necessary adjustments to ensure continuous effectiveness. Any tax revenue should be reinvested in educational and informational programs on the health risks of prolonged engagement with those internet activities, providing free mental health care and counseling for at-risk subgroups, and funding research in this area by independent third parties.

### 4.1 Pros and Cons of the Proposed Approach

As with every policy instrument, the proposed policy approach is not perfect. In the paragraphs that follow, we consider and respond to some of the main critiques.

**Critique 1**—“Calculating the tax imposes a heavy computational and privacy burden.”

It can be argued that the proposed policy instrument requires a heavy lift on the regulated company (and the regulatory entity) in computing the tax since it would require data on the individual use of each online platform, across multiple devices and across time. Along the same lines, sharing individual data on time spent on an online platform might be intrusive to personal privacy.

Nevertheless, individual data on time spent on an online platform represent information that most of these platforms already collect and use. The computation of the tax is a simple elaboration of those data, simpler than most elaborations that these companies might already conduct. Since computing the tax requires only individual time spent on the online platform devoid of any additional information (like precise location, type of device, age, etc.), the data can be effectively anonymized if they have to be shared with a regulator for monitoring, thus limiting any possible impingement on individual privacy.

**Critique 2**—“Companies will shift the burden of the tax onto users.”

A common pushback to a tax instrument is that if users perceive the taxed item to be a necessary part of their daily lives and if they cannot find any substitute, then companies might shift the burden of the tax onto consumers by increasing the prices of the product or service they provide. When products are addictive, then consumers are particularly vulnerable because they will have a hard time curbing consumption.

While this is a major critique of the taxation of addictive products like tobacco, the case of online platforms is substantially different. As many internet memes suggest: users are not the consumers; they are the product. Many of the platforms under consideration in this paper (video games, social media, and online pornography) offer standard functionalities that are free to users (or charge very modest fees to incentivize wide uptake).

Therefore, the technology companies that own the regulated platforms could pass the tax only onto their real clients: companies that purchase advertisement space. So we would expect that the cost of advertisement on these platforms might increase following the introduction of a tax and that the tax burden would partially be shifted from the regulated technology companies to the companies who purchase advertisement space.

This is far from a negative or unintended effect of this system. Online platforms are generating advertisement space whose market is producing a negative social externality in the form of a negative mental health impact on users. This negative externality is unaccounted for within the market for online advertisement, and the proposed tax generates a system to make sure that buyers and sellers of advertisement space incorporate that social cost into their transactions.

**Critique 3**—“This system will just lead to a flourishing of various online platforms with users hopping from one to the other.”

One of the risks of the proposed system is that technology...
companies will create multiple online platforms and encourage users to hop from one to the other in order to evade the time limit. Here it is good to remember that the proposed regulation is supposed to limit technology companies and not their users. This is the reason why it is essential that if such a system is implemented, the final regulation will need to introduce language that forbids the technology companies from altering the functionalities of the online platform for those users who are approaching or have surpassed the set time limit.

Nevertheless, it is still possible that ultimately the system is imperfect and therefore that there might be some increase in the number of platforms available to users. This is not necessarily a bad outcome, however. Behavioral research shows that introducing natural stopping points when using a digital tool decreases the time that users spend on the tool. This is why online platforms strive to achieve a smooth design devoid of stopping points (infinite scrolling and autoplay features are good examples). Having to switch to a different online platform would likely create a natural stopping point in the way users experience the platform and therefore might create some inertia and decrease overall use.

**Critique 4**—“This system will impose a similar restriction to very heterogeneous companies.”

A major component of the design of this tax system is setting the time limit \( x \). Suggesting precise values for those time limits is out of the scope of this paper and is left to future research. It might be worth considering different values for different industries. For example, the time limit for video games should be different from the time limit set for social media, since the former appears to be more addictive than the latter. Nevertheless, this distinction between the two might become more blurred if online platforms become hybrids of different products, as in the case of Facebook offering in-application games such as Farmville. This might be an argument in favor of a homogeneous time limit.

**5. Conclusions**

This policy solution is not a panacea, but it might provide a sufficient incentive for companies to discourage compulsive behavior, or at least it should provide the public with the resources to respond to the mental health damages that co-occur with prolonged interactions on internet platforms. Given the wealth of user data collected, calculating this tax should be both feasible and not burdensome on the part of the regulated companies. Since the only data that would be shared with the government in case of verification are the times spent on the platform by individual users, it will not impact user privacy.

Finally, the tax would not impact healthy users but might greatly benefit those who are more vulnerable to the potential side effects of prolonged engagement with internet gaming, social media, and online pornography. The evidence collected so far suggests that only continuous and potentially compulsive use of these digital technologies is associated with the deterioration of mental health. Therefore, this policy intervention should affect only behavior with negative health consequences, without impacting other users.

Most importantly, the suggested policy is not meant to be enacted in isolation. Other responses, such as continued advocacy to destigmatize these mental health issues, providing freely available mental health resources and informational programs on the health risks, and advancing the scientific research on these issues, are all important parts of responding to this mental health crisis. Digital technologies will continue to improve their ability to capture user attention and emotionally manipulate users. It is important to proactively think of ways to regulate these industries to prevent those mental health disorders from becoming the norm rather than the exception.
References


