



## ■ artículo

SCV Societat Catalana  
de Victimologia

SOCIEDAD VASCA DE VICTIMOLOGÍA  
SOCIAL ESTEREOLOGIA SOCIETATIS

HUYGENS  
EDITORIAL

REVISTA DE VICTIMOLOGÍA | JOURNAL OF VICTIMOLOGY  
Online ISSN 2385-779X  
www.revistadevictimologia.com | www.journalofvictimology.com  
DOI 10.12827/RVJV.19.02 | N. 19/2025 | P. 43-76  
Fecha de recepción: 01/01/2024 | Fecha de aceptación: 01/12/2024

# Persuasión algorítmica y adicción tecnológica: las nuevas tabaquerías<sup>1</sup>

M<sup>a</sup> Isabel González Tapia

Universidad de Córdoba; fd1gotam@gmail.com

## Resumen

El objeto de este trabajo es cuestionarse si el Derecho Penal habría de intervenir o no en el control de los riesgos derivados del diseño adictivo de las redes sociales y plataformas, a la vista de las conclusiones aportadas por la literatura científica especializada. Se concluye afirmativamente sobre esta cuestión, sin perjuicio de la tutela administrativa que pueda también darse, formulando los rudimentos de una propuesta político-criminal incriminadora, articulada en dos pilares fundamentales. El primero de ellos es la preponderancia de dos principios fundamentales: el criterio de precaución, frente a riesgos muy importantes o sistémicos, y el principio de responsabilidad *pro víctima*, que prioriza el control del riesgo ante la posibilidad de los daños y riesgos que la literatura científica conecta al diseño adictivo de las redes sociales y plataformas. Como segundo pilar, considerar estos diseños adictivos fuera del riesgo permitido con relación a los menores, sirviéndonos del *paragüas* que proporciona ya el art. 5 del Reglamento IA de la UE y centrándonos muy particularmente, en la responsabilidad de las plataformas que usan tales herramientas de IA para conseguir la atención y la constante y compulsiva conexión del usuario.

## Palabras clave

Adicción a redes sociales, IA persuasiva, algoritmos adictivos, responsabilidad penal por diseño algorítmico adictivo

## Abstract

The purpose of this work is to examine whether Criminal Law should intervene in the regulation of risks associated with the addictive design of social networks and platforms, in light of the findings provided by specialized scientific literature. This analysis reaches an affirmative conclusion on the matter, without prejudice to the administrative oversight that may also apply, and outlines the foundational elements of a criminal policy framework based on two fundamental pillars. The

1 Este trabajo se enmarca en el Proyecto de Excelencia de la Junta de Andalucía “Criminalidad en contextos digitales de ocio desviado: Alternativas posibles contra una economía de consumo deshumanizada”, Universidad de Sevilla, (Ref. PROYEXCEL\_00621).



first pillar emphasizes two key principles: the precautionary principle, which addresses significant or systemic risks, and the pro-victim responsibility principle, which prioritizes risk control to prevent potential harms and dangers linked by scientific literature to the addictive design of social networks and platforms. The second pillar involves classifying such addictive designs as falling outside the scope of permitted risk concerning minors. This approach leverages the protective framework already established by Article 5 of the EU AI Regulation, focusing specifically on the liability of platforms that employ such AI tools to capture users' attention and foster constant and compulsive engagement.

### Keywords

Social network addiction, persuasive AI, addictive algorithms, criminal liability for addictive algorithmic design

## 1. Introducción

En los últimos tiempos se ha venido advirtiendo de la aparición de nuevas adicciones relacionadas con las nuevas tecnologías de la sociedad de la información o internet. Más reciente todavía, y de forma aún más alarmante, por aquellas herramientas mediadas por la Inteligencia artificial (en adelante IA). En general, todas estas adicciones se agrupan en torno al concepto de “adicciones digitales” o “tecnológicas” y se englobarían, a su vez, dentro del grupo de “adicciones sin sustancia” o “comportamentales”(Ding et al., 2023; OEDA, 2023).

Por ejemplo, se han reportado multitud de casos de adicción al juego online y a los videojuegos, que emplean algoritmos que adaptan el juego al comportamiento del usuario, lo que puede fomentar una mayor inmersión, haciendo que los jugadores se sientan constantemente desafiados o recompensados. En casos extremos, esta manipulación algorítmica puede llevar a la adicción, en la que los individuos pierden el control sobre el tiempo que pasan jugando, afectando su vida personal, profesional y a su salud mental(King & Delfabbro, 2020; Delic et al., 2023). La adicción a los videojuegos, denominada como *Trastorno por uso de videojuegos* está reconocida en el manual de *Clasificación Internacional de Enfermedades* (en adelante, CIE 11), Código: 6C50, distinguiéndose si es predominantemente en línea o no. También en el *Manual Diagnóstico y Estadístico de los Trastornos Mentales* (en adelante DSM-V-TR) se ha reconocido en 2023 el *Internet Gaming Disorder*, en la sección de condiciones recomendadas para futuras investigaciones. Así mismo, se han descrito múltiples casos de adicción a las redes sociales, que han estado en el origen de devastadores daños personales entre niños y adolescentes en todo el mundo desarrollado (Cfr. p.e. Costello et al., 2024). Las redes sociales, apoyadas en IA persuasiva, están diseñadas para captar y mantener la atención de los usuarios, apoyándose en un diseño adictivo que puede hacer que los más jóvenes desarrollen dependencia y múltiples problemas asociados a esta circunstancia. A este problema, dentro del grupo de adicciones tecnológicas, nos vamos a referir



específicamente en este trabajo. Todavía de forma más reciente, se ha comenzado también a describir casos de adicción y de manipulación a través de IA generativa, chatbots, robots y asistentes virtuales (Matz et al., 2024). Las personas pueden crear vínculos emocionales con estas tecnologías y, en casos extremos, dicha interacción constante, con IAs que imitan el comportamiento y el lenguaje humano, puede generar *dependencia emocional*. Algunos individuos pueden llegar, incluso, a considerar estas IAs como confidentes o fuentes de compañía, lo que afecta sus relaciones interpersonales en el mundo real y los convierte en sujetos fácilmente manipulables. Además de esta dependencia emocional, la literatura científica señala otros riesgos vinculados con la IA generativa relacionados con la dependencia en la *creación de contenido*, que pueden afectar a la capacidad crítica y creativa del ser humano (S. Zhang et al., 2024).

El avance de la IA y la interacción humana con ella plantea múltiples y radicales desafíos legales y éticos. La adicción digital es sólo uno de ellos, pero de primera magnitud. El aumento de casos de adicciones digitales y las consecuencias negativas que pueden derivarse para el ser humano, particularmente para los más jóvenes, hace surgir la pregunta de si el derecho penal debería o no desempeñar un papel más activo en la protección de estos usuarios.

Tal vez, como cuestión previa, podríamos preguntarnos acerca de la incidencia real de esta problemática. Puede anticiparse que, aunque los datos que arroja la literatura científica no puedan considerarse absolutamente concluyentes, por la relevancia de otros factores, lo cierto es que hay múltiples fuentes y un vasto cuerpo doctrinal que apuntan a la importancia de estos riesgos y a su mayor prevalencia entre los adolescentes del mundo entero. En el trascendental informe emitido por The U.S. Surgeon General's Advisory (2023) se ha llegado a calificar el problema de la adicción y uso compulsivo de las redes sociales entre los niños y adolescentes, como una *epidemia de salud mental*. Así, se señala, entre otros datos, que las muertes relacionadas con la salud mental han aumentado en un 57% y que también se observa un significativo incremento de los pensamientos suicidas y en el reconocimiento de la adicción a las redes sociales. Se reporta, además, la ambivalencia que sienten los adolescentes con su uso, puesto que no necesariamente disfrutaban de la experiencia. En este informe también reconoce, sin embargo, que la afectación en la salud mental del uso de las redes sociales depende de múltiples factores, que justificarían significativas diferencias individuales, como el tiempo que se pasa en red, el tipo de contenido que se consume, las interacciones que se realizan y las diferencias personales entre los distintos individuos, en su naturaleza, contexto socioeconómico y cultural... (Vaid et al., 2024; W. Zhang et al., 2022).

Similares conclusiones se han alcanzado en otros informes especializados. El *Center for countering Digital Hate* ha reportado otros datos que no pueden



por menos que hacernos pensar. El algoritmo de TikTok se dirige a los adolescentes con contenido de imagen corporal y de salud mental cada 39 segundos, con contenido relacionado con el suicidio cada 2 minutos y 36 segundos, y con contenido relacionado con trastornos alimentarios cada 8 minutos... Tal oferta de contenidos (y en ese formato completamente invasivo) podría exacerbar los problemas de salud mental entre los espectadores impresionables. Además, para evitar que se puedan hacer estimaciones, está eliminando el contador de sus hashtags (CCDH, 2024). El informe del *Pew Research Center para EEUU* (Anderson et al., 2023) señala que el uso de las redes sociales entre la adolescencia puede considerarse universal: por encima del 93% de los jóvenes entre los 13 y los 17 años usan alguna plataforma prácticamente a diario y un quinto afirma usarlo de forma constante, siendo especialmente prevalente el uso constante por parte de las chicas adolescentes. Prácticamente desde 2014-15, el uso constante de las redes sociales se ha duplicado entre los jóvenes, siendo YouTube y entre los más jóvenes TikTok, Snapchat e Instagram las redes que son más utilizadas. Además, el uso de redes sociales entre niños/as de 8 a 12 años se situaría en torno al 40% (Rideout, 2021). Aproximadamente un 11% de los adolescentes estarían presentando síntomas de uso patológico o de adicción a las redes sociales: falta de capacidad de control del uso, abstinencia, depresión y ansiedad, abandono de otras actividades o responsabilidad... (The Lancet, 2024). Además, es consistente la afirmación de que el uso abusivo de las redes sociales tendría un especial impacto en la salud mental de las niñas y chicas adolescentes, un impacto dos veces mayor en sentimiento de soledad y en riesgo de sufrir trastornos alimentarios y problemas académicos por falta de concentración derivada del diseño adictivo de redes. (GEM UNESCO, 2024)

En España, el *Informe sobre Trastornos Comportamentales de 2023* (OEDA, 2023) ha recoge que el uso de internet entre la población española es universal (entre 15 y 65 años), con un uso lúdico medio que varía y que en la población de entre 14 a 18 años es de 5.36horas/entre semana y 6.97 horas/fin de semana. Además, señala que, al igual que ocurre con la población general (Savolainen & Oksanen, 2024), los consumos intensivos de drogas y alcohol son mayores entre los estudiantes con posible uso problemático de internet. En cuanto a la prevalencia de usuarios de internet con un posible uso problemático, entre la población de 15 a 24 años sería del 10,8% en 2022, frente al 5.3 entre la población de 25 a 34 años, señalando igualmente el impacto mayor de esta problemática entre las mujeres y niñas adolescentes. En cuanto a las admisiones a tratamiento por adicción (a 2021), la que mayor genera es la adicción al juego (51% en mujeres y 85% en hombres), mientras que la adicción o uso compulsivo a internet, móviles o dispositivos electrónicos, redes sociales o videojuegos sería del 16,5 % en mujeres y el 9,2% en hombres. El estudio de (ONTSI, 2024) sobre el uso de las tecnologías por los menores en España, con datos de 2023, señala que el



uso de internet entre los menores es prácticamente universal, con un uso por encima del 94% entre los niños/as de 10 y 15 años y con una disponibilidad de teléfono móvil en el 70.6%. De este mismo organismo, en el informe sobre el *Impacto del aumento del uso de internet en las redes sociales en la salud mental de jóvenes y adolescentes* (ONTSI, 2023) se extraen los siguientes datos: en cuanto al riesgo de *uso compulsivo*, se detectó que el 11,3% de los jóvenes entre 15 y 24 años, pero un 33% entre los adolescentes de 12 a 16 años, presentan un riesgo considerable de uso compulsivo de redes sociales. Con relación a la *ideación suicida*, en 2021, el 9,6% de los adolescentes y jóvenes manifestaron pensamientos suicidas, un dato que UNICEF eleva hasta el 10,8%. Este aumento es preocupante y muestra la conexión entre el uso excesivo de redes y los problemas de salud mental. En cuanto al *estrés*: el 21,1% de los jóvenes considera que la interacción constante en redes les genera niveles elevados de estrés, exacerbando los problemas emocionales derivados del uso de tecnología; y en cuanto al *impacto académico*: el 44,6% de los estudiantes afirman que el tiempo invertido en el uso de tecnologías les quita horas que podrían dedicar al estudio, lo que afecta negativamente su rendimiento académico. Además, el estudio también subraya las *desigualdades de género* en estos impactos, señalando que un tercio de las adolescentes estarían afectadas de manera desproporcionada. En el estudio epidemiológico de (Nogueira-López et al., 2023), sobre el uso problemático de Internet y el juego online entre los adolescentes españoles, presenta una prevalencia diferente según la medición responda a los criterios conectados con la descripción del DSM-V o del CIE 11. Su conclusión final sería que la utilización de los criterios de referencia del DSM-V implicaría una sobrepatologización de esta problemática, frente a la visión más realista del CIE-11. Así, utilizando una medición inspirada en el marco DSM-5, han hallado una prevalencia del 33% para el uso problemático de Internet y del 3,1% para el juego. Con un enfoque más conservador, inspirado en el CIE-11, las tasas de prevalencia serían del 2,98% para el uso problemático de Internet y el 1,8% para el juego problemático. Con relación a los factores de riesgo relacionados, la varianza se relaciona con el género. En el caso del femenino, serían predictores la educación superior de los padres, el tiempo elevado de conexión a Internet, el reporte estar en línea después de la medianoche, así como el uso del teléfono móvil en clase. En el masculino, con relación al juego, serían relevantes la “situación de vida” en la que las familias no tienen una estructura tradicional o un entorno estable, el tiempo elevado de conexión a Internet y el uso de móvil en clase.

Se considera que, siquiera una minoría de los jóvenes expuestos a las redes sociales, podrían convertirse en adictos a las mismas, en los mismos términos que experimentarían en una adicción a sustancias o al juego (Kuss & Griffiths, 2017; Turel et al., 2014a), variando la prevalencia de la adicción general a internet entre 12.6% al 67.5%, por la relevancia de las diferencias intercultural-



les (Kuss et al., 2021) this review summarizes and investigates studies published outside of Europe between April 2013 and September 2019. Literature on Internet addiction was included and reviewed critically following the PRISMA guidelines, with the objectives to (i. Griffiths (2018) señala que, aunque el número de adolescentes que pudieran ser considerados “adictos”, en sentido estricto, pueda no ser considerado elevado, no puede olvidarse que muchos otros considerados “heavy users” o usuarios habituales de estas redes sociales pueden también experimentar consecuencias problemáticas derivadas relacionadas con las herramientas de “enganche” psicológico se usan en estas redes sociales y plataformas.

A la vista de estos datos, el riesgo de adicción a las redes sociales y los potenciales daños asociados merece, cuando menos, una consideración detenida por parte del Derecho Penal. Y desde esta perspectiva, este trabajo pretende ser, al menos, una primera aproximación político-criminal a esta problemática.

## 2. Riesgos y daños descritos con relación al uso abusivo o compulsivo de las redes sociales

El diseño adictivo de las plataformas y redes sociales sería, de una parte, el vehículo catalizador del potencial daño derivada de la sobre-exposición al contenido lesivo y también sería, de otra, agente causante de daño a la salud mental de las personas, provocado por la adicción generada.

### 2.1. Creación o incremento del riesgo para la salud mental, derivados de la sobre-exposición a contenidos lesivos y/o manipulativos.

Aparte de la exposición a una publicidad constante y muchas veces encubierta, los menores están expuestos a contenidos altamente inapropiados, de fácil e incontrolado acceso, y susceptibles de incentivar comportamientos antisociales y autolesivos. El contenido inapropiado y lesivo al que están expuestos los menores abarca desde mensajes sexistas, violentos, de odio, pornográficos, hasta aquellos que fomentan las autolesiones y el suicidio. Estos contenidos pueden ser buscados activamente, encontrados accidentalmente o sugeridos algorítmicamente. El uso constante de estas redes incrementa los riesgos derivados de la exposición a tales contenidos lesivos, y un uso compulsivo los exacerbaría aún más, especialmente entre menores que ya presentan vulnerabilidades individuales en su salud mental (Sumner et al., 2021). La trágica historia de Molly Russell, por ejemplo, ilustraría la relación entre la exposición a contenido dañino en redes como Instagram y Pinterest y el suicidio. Molly, una adolescente británica de 14 años, estuvo expuesta repetidamente durante meses a material sobre au-



tolesiones y suicidio, lo que agravó su depresión y finalmente contribuyó a su muerte (Dyer, 2022). Y no ha sido un caso aislado (Bloomberg.com, 2022).

Otra problemática reportada y bien establecida en la literatura es la relacionada con el desarrollo o incremento de dismorfia corporal y de trastornos alimentarios, que se conectaría con el consumo de contenidos específicos de fomento de tales comportamientos, pero también de la propia dinámica de las plataformas basadas en la imagen, como p.e. Instagram o TikTok. Estas redes emplean inteligencia artificial para personalizar el contenido que muestran al usuario y transmiten incansables imágenes de ideales de belleza y de perfección, especialmente impactantes entre los usuarios adolescentes. Esto puede conducir a una obsesión con la imagen corporal, especialmente entre los y, sobre todo, las jóvenes, que se comparan permanentemente con figuras de belleza no realistas; puede conducir, pues, a fomentar conductas de riesgo para trastornos alimentarios como la anorexia o la bulimia. La dismorfia corporal, que implica una percepción distorsionada de la propia apariencia, es una consecuencia cada vez más visible en jóvenes expuestos a estos estándares (e.g. Lonergan et al., 2020; Lozano-Muñoz et al., 2022) entre otros muchos.

Además, puede afirmarse también que existe otra correlación bien establecida en la literatura científica entre el uso constante y abusivo de las redes sociales y el desarrollo o incremento de ansiedad y depresión. La comparación continua con los logros y la vida de otros, común en plataformas como Instagram o Facebook, provoca sensaciones de insuficiencia y la idea de que los demás se divierten más, tienen una vida mejor y son más valiosos socialmente. Esto puede desarrollar o agudizar baja autoestima y desencadenar trastornos de ansiedad y depresión, particularmente entre las chicas adolescentes (Cfr. en W. Chen & Lee, 2013; Vogel et al., 2014; Elhai et al., 2019; Lee et al., 2021; Şenol et al., 2024; Andreassen et al., 2017; Shannon et al., 2022...). Así mismo, Ryding et al., (2024), han precisado que tales efectos negativos se vincularían sobre todo al uso problemático *pasivo* de las redes sociales, entre otros muchos.

## 2.2. Riesgos específicamente relacionados con el diseño adictivo de las plataformas y redes sociales

Se ha señalado que el uso abusivo de las tecnologías digitales puede provocar deterioro cognitivo en diferentes vertientes (Cfr. Shanmugasundaram & Tamilarasu, 2023) y, más en concreto, con relación a las redes sociales y plataformas similares, deterioro por fragmentación de la atención y desconcentración (Firth et al., 2019). Las redes sociales determinan al usuario a enfocar su atención en diferentes objetos e informaciones por un espacio muy corto de tiempo (por ejemplo, un mensaje de TikTok que dura de media 15 segundos), diariamente, durante horas al día y desde la infancia. Con este diseño adictivo, como ha



afirmado (Dino Ambrosi, 2023), se induce a que el propio usuario se entrene a sí mismo para convertirse en un *sujeto patológica o crónicamente distraído*. Este bombardeo constante de estímulos en las redes sociales afecta la capacidad de mantener la atención (Y. Chen et al., 2023; Farchakh et al., 2020), lo que repercute también en el rendimiento académico y en otras actividades que requieren concentración sostenida. Dicha problemática se vería maximizada con el uso de la IA, que puede crear permanentemente contenido personalizado a partir de *bots* o usuarios ficticios (UNESCO, 2024). Al mismo tiempo, la omnipresencia de la idea de “lo breve”, “lo efímero”, “el slogan”, la “burbuja” que se crea a través de contenidos sugeridos, la “banalización” de lo importante... van reduciendo la capacidad de tomarse el tiempo suficiente para comprender fenómenos complejos, para razonar, para considerar la opiniones y argumentos de los otros... contribuyendo a un empobrecimiento cognitivo cada vez más evidente, a la radicalización, a la polarización de las opiniones...

Además, el uso compulsivo de estas redes y plataformas, exponen al usuario a un mayor riesgo de para su privacidad e imagen, por la difusión de imágenes íntimas y la exposición digital excesiva. La falta de regulación y supervisión en las redes sociales expone, particularmente a los menores, a peligros como el de la difusión de imágenes íntimas para cumplir “retos”, o inducidas para crear un contenido “más impactante”, que la IA ayuda a viralizar, lo que puede llevar a situaciones extremadamente dañinas, como el ciberacoso, depresión e incluso la autoflagelación o el suicidio. Este conocimiento es, desgraciadamente, de dominio público.

De otra parte, aunque las redes conectan virtualmente a los usuarios, muchas veces reemplazan las interacciones físicas y contribuyen al aislamiento en el mundo real, disminuyendo las habilidades sociales y afectando la creación de relaciones auténticas. Los usuarios podrían depender de *chatbots* o asistentes virtuales para compañía emocional, lo que podría generar aislamiento social y agravar problemas mentales. Así mismo, el tiempo excesivo dedicado a redes sociales o a la IA para entretenimiento o apoyo emocional lleva a la procrastinación y al abandono de obligaciones académicas, familiares o laborales, disminuyendo la productividad y generando tensiones interpersonales. Podría añadirse, aunque no pueda considerarse como un daño a la salud mental, propiamente dicho, otra consecuencia derivada del diseño persuasivo algorítmico en las redes sociales como es el riesgo de manipulación del usuario y de la sociedad en su conjunto. Hace tiempo ya que (Jaron Lanier, 2018) advertía que, probablemente, el principal riesgo de las redes sociales es el de la manipulación del comportamiento del ser humano, dirigiéndolo hacia lo que las plataformas o sus clientes quieren conseguir: estados de opinión, manipulación del voto en unas elecciones, compras de determinados productos. Las plataformas y redes sociales se ocuparían “de la adicción”, mientras que los clientes se ocuparían de



“la manipulación” de los usuarios con sus mensajes, y sometiéndolos, como a la sociedad, a una versión atenuada pero constante de modificación del comportamiento.

### 3. Adicción o uso compulsivo a las redes sociales y otras herramientas mediadas por la IA persuasiva

#### 3.1. Diseño adictivo de las plataformas y redes sociales: mecanismos psicológicos de la adicción tecnológica

IA persuasiva utiliza mecanismos psicológicos para fomentar el uso compulsivo de las tecnologías y su presupuesto, ello es evidente, es la información sobre el usuario, que permite saber, a través de un monitoreo constante, lo que le gusta, lo que le llama la atención, sus temas y actividades de interés, en lo que gasta su dinero, lo que le hace reaccionar... Y se lo ofrece algorítmicamente, a través de contenido sugerido personalizado sin fin, para mantener la atención del usuario y conocerlo cada vez mejor (Keipi et al., 2016). El mecanismo es, pues: el monitoreo constante del usuario, la modificación progresiva, discreta y calculada del contenido que se le ofrece y la utilización de diferentes herramientas adictivas, que se han incluido intencionadamente, para incrementar el tiempo de atención (Singer, 2015). Se perseguiría, describe Jaron Lanier (2018), algo así como una *adicción encubierta y estadística*, incitando al conjunto social a usar las redes sociales al máximo posible. En el proceso, utilizan herramientas y diseños dirigidos a generar adicción, activando el sistema neural de recompensa, y mantener así el tiempo que el usuario está conectado y su atención (Montag et al., 2019). Y a partir de ahí, poder introducir soterradamente pequeños cambios dirigidos, hacia donde interesa al cliente de estas empresas (Matz et al., 2017). Sería, como han reconocido los propios ingenieros de estas empresas, un deliberado diseño adictivo, que se inserta en un modelo de la economía de la vigilancia (Zuboff, 2019) y de la atención (p.e. Alter, 2017; Wu, 2016). Y a partir de ahí también, se generaría la adicción a las actividades que los usuarios realizan constante y compulsivamente en internet: sexo, compras, juegos online, juegos de azar, activismo social y político, posteo de la vida privada..., más aún cuando se engarza con otro elemento de pertenencia y cohesión como son las comunidades online (Oksanen et al., 2024).

Resulta pertinente, por tanto, preguntarse acerca de cuáles son los mecanismos y herramientas básicas utilizadas en el diseño adictivo de las plataformas o redes sociales. En síntesis, podría decirse que se basan en la personalización, la utilización de recompensas variables intermitentes, en la explotación de necesidades psicológicas y en la gamificación.



### *Monitoreo constante del usuario y contenidos dirigidos algorítmicamente*

Entre los mecanismos utilizados en el diseño adictivo se encuentra, como ya se ha señalado, la personalización de los contenidos dirigidos algorítmicamente. La inteligencia artificial monitorea el comportamiento del usuario constantemente, para ofrecer contenido específico de interés, que aumenta la probabilidad de interacción y permanencia en la plataforma, creando una “*burbuja*” de contenido *personalizada*. El diseño del “*scroll infinito*” crea un flujo continuo de contenido sugerido, alentando a los usuarios a deslizarse sin fin (p.e. el “Para ti” de X o la “Reproducción automática” de Youtube...), puesto que, no debe olvidarse, el algoritmo está permanentemente ajustándose para ofrecer contenido personalizado que dificulte la desconexión. Este mecanismo induce una experiencia similar a un juego, donde la acción de desplazarse se vuelve inmersiva y difícil de interrumpir, uniéndosele a todo ello el que no hay limitaciones temporales en el uso de la aplicación, ni muchas veces referencias temporales a cuando el contenido fue subido a la red o del tiempo que el usuario está utilizando el servicio. La idea es que, cuanto más tiempo permanezca el usuario, mejor; esa es precisamente la clave y la finalidad.

### *Explotación de la novedad y de la inseguridad*

La explotación de la inclinación hacia la novedad y de la inseguridad del ser humano es otro mecanismo utilizado en el diseño adictivo, a través de notificaciones incesantes y de contenido efímero. La explotación de estas vulnerabilidades se ha conectado con el “miedo a perderse algo” o FOMO (“*Fear of Missing Out*”), que se refiere al temor persistente de que los demás estén disfrutando experiencias valiosas o gratificantes y el usuario no, lo que genera una ansiedad social, por sentir que “no se está al tanto” o “que no se participa” en los eventos que se visualizan o en las conversaciones que ocurren en el entorno digital (Oberst et al., 2017) but it has also been shown that these users can suffer from negative psychological consequences when using these sites excessively. We analyze the role of fear of missing out (FOMO. Las plataformas de redes sociales están diseñadas para amplificar este miedo a través de mecanismos como las *notificaciones incesantes*, que constantemente alertan a los usuarios de nuevas interacciones, publicaciones o mensajes que podrían estar perdiéndose, si no revisan sus dispositivos de manera frecuente. Estas notificaciones interrumpen el día a día de las personas, creando una necesidad de revisar continuamente la plataforma para no quedar excluidos de lo que está sucediendo. Otro factor clave en la amplificación del FOMO es la *naturaleza efímera* del contenido, que se presenta en plataformas como *Snapchat*, *Instagram Stories* y otras aplicaciones que promueven la creación de publicaciones que desaparecen en un plazo corto (generalmente 24 horas) o son meros *directos*. La temporalidad de este contenido



genera una sensación de urgencia en los usuarios, que, no solo impulsa a los usuarios a consumir el contenido rápidamente, sino también a crear y compartir sus propias experiencias en tiempo real, antes de que desaparezcan o queden “obsoletas”, reforzando un estado de alerta y ansiedad constantes en el usuario. Además, el FOMO fomenta una comparación social constante y que los usuarios sientan que deben estar al mismo nivel de las experiencias y actividades que observan en los perfiles de los demás (Cfr. Soriano-Sánchez, 2022, por todos).

### *Explotación de la necesidad de aprobación o validación social*

Otro de los ingredientes clave que facilita el uso habitual e intensivo de las redes sociales, especialmente entre adolescentes, es el botón de “me gusta”. Esta función, introducida por Facebook en 2009, ha resultado ser un mecanismo sumamente efectivo para captar la atención de los jóvenes de manera constante (Sherman et al., 2016). Aunque su diseño es simple, explota una necesidad humana profunda de validación social, que es particularmente poderosa durante la adolescencia. En esta etapa de la vida, los individuos están en un proceso de formación de su identidad, y la búsqueda de aprobación y pertenencia es crucial. Los “likes” funcionan como una métrica visible de aceptación social, lo que hace que los usuarios, especialmente los adolescentes, vuelvan una y otra vez a buscar este tipo de recompensa. Diversos estudios han señalado cómo este tipo de características refuerzan el “ansia de validación” en los adolescentes y suponen una especie de “secuestro” del sistema de recompensa del cerebro del usuario, que se ve impulsado a volver constantemente a las plataformas, buscando la próxima *dosis* de aprobación social, como si estuviera pendiente de una máquina expendedora de reconocimiento (Cfr. Morgans, 2017) , por todos.

Plataformas como Instagram, por ejemplo, inducen a los usuarios a participar en dos conductas que se retroalimentan: la monitorización constante de sus vidas y la necesidad de plasmarla en una fotografía, video corto o “story” de contenido efímero. Después de publicar, el usuario entra en un estado de espera, *ansioso por los “likes”* que validarán su publicación, perpetuando esta dinámica de dependencia emocional. El uso de *filtros* y la participación en “*challenges*” también forma parte de esta explotación de la necesidad de validación. Los filtros no solo embellecen la realidad, sino que se convierten en una herramienta para conformarse a los estándares de belleza y comportamiento socialmente aceptados, lo que aumenta la presión social. Los *challenges virales* incentivan la participación en tendencias, donde los adolescentes no solo buscan aprobación, sino también pertenecer a una comunidad, validando su lugar en el grupo. La obtención de “likes” y otras métricas (como comentarios o vistas) se convierte en una *medida de autoestima y de valor personal*, lo que genera una constante búsqueda de aceptación y reconocimiento (Nesi & Prinstein, 2015).



No es solo el potentísimo significado del “*botón like*” u otro tipo de refuerzos, sino el modo en que se presentan e interactúan con el usuario; particularmente adolescente. Esta dinámica adictiva explota la necesidad de validación social y afecta directamente el sistema neurológico de la recompensa, como en el juego por internet (Turel et al., 2014b), sirviéndose de herramientas como las recompensas variables intermitentes (RVI) y los “trofeos” al uso extremo.

Las RVI, como su nombre indica, son un refuerzo positivo intermitente con la particularidad de que su aparición no es predecible por el usuario. Esto es esencial para el éxito de la adicción: cuando no sabemos cuándo llegará la próxima recompensa, estamos más inclinados a continuar buscando, manteniéndonos más tiempo conectados. Es una estrategia de gamificación que ha demostrado ser eficaz en juegos y plataformas de entretenimiento. Se ha convertido en un elemento clave en la estrategia adictiva de las redes sociales y plataformas. Los usuarios reciben contenido o recompensas en momentos estratégicos para maximizar su tiempo de uso y evitar que se desconecten (Cfr. Clark & Zack, 2023)2023. Este sistema de recompensas no se entrega de forma constante o predecible, sino de manera aleatoria o esporádica, lo que genera en los usuarios un alto nivel de expectación y ansiedad por el próximo “like”, ya que no saben cuándo ni cómo recibirán la siguiente recompensa (Griffiths, 2018). Además, se ha sugerido que el breve retraso que ocurre al abrir una aplicación de redes sociales es una estrategia intencional para aumentar la sensación de anticipación, ya que, psicológicamente, la espera de una recompensa puede liberar casi tanta dopamina como la recompensa misma (Morgans, 2017).

Finalmente, los “trofeos” al uso extremo son otra herramienta clave. Las plataformas premian a los usuarios más activos con insignias, niveles o recompensas simbólicas, creando un sentido de logro y validación. Aunque estas recompensas son superficiales, generan una sensación de progreso y éxito que empuja a los usuarios a continuar. Un usuario que ha recibido múltiples trofeos se sentirá más inclinado a seguir utilizando la plataforma para no “perder” ese progreso, perpetuando la adicción. Un ejemplo claro de esto es la acumulación de “rachas” o el seguimiento del número de días consecutivos en los que el usuario ha interactuado con la plataforma, creando presión psicológica para seguir utilizando la aplicación sin interrupciones (Carr & Rosaen, 2024).

### 3.2. Especial vulnerabilidad de menores y adolescentes. ¿Por qué este diseño adictivo es tan dañino para niños y adolescentes?

El mayor impacto de la adicción a las redes sociales y su lesividad asociada en niños y adolescentes se debe a una combinación de factores relacionados con el desarrollo cognitivo, emocional y social de este grupo poblacional, así como con las estrategias de diseño específicas implementadas por las plataformas (The



U.S. Surgeon General's Advisory, 2023; Abrams, Zara, 2024; Lopez-Fernandez et al., 2024; Huang et al., 2024; Gingras et al., 2024 and few studies have considered the moderating role of extraversion. In 2017 (T1, entre otros).

Una serie de factores clave explicarían por qué este diseño es tan lesivo para los menores:

*Por las vulnerabilidades derivadas del desarrollo cerebral en formación:*

Durante la adolescencia, el cerebro está en proceso de maduración, presentándose una “segunda ventana de oportunidad” al ser humano para conformar la personalidad futura de la persona. En dicho periodo, existe un bajo control de los impulsos, una elevada asunción de riesgo y una especial sensibilidad al daño y al estrés social (Fuhrmann et al., 2015; Romer, 2010; Murimi, 2023, entre otros muchos). La especial vulnerabilidad de niños y adolescentes frente a la adicción a las redes sociales y los daños derivados está intrínsecamente vinculada a la circunstancia de que las herramientas adictivas usadas inciden y afectan al sistema de recompensa y manipulan la liberación de dopamina en sus cerebros en desarrollo, que tienen una mayor sensibilidad afectiva y social (Towner et al., 2023). Así, la dopamina es un neurotransmisor asociado con la recompensa y el placer. El código diseñado en las plataformas de redes sociales está específicamente orientado a manipular la liberación de dopamina en los cerebros en desarrollo de niños y adolescentes. La anticipación de recompensas, como obtener “Me gusta” o participar en desafíos, activa el sistema de recompensa del cerebro, generando una sensación placentera. Entre los 10 y los 12 años, los cambios en el cerebro hacen que las recompensas sociales comiencen a sentirse mucho más satisfactorias. Específicamente, los receptores de las “hormonas de la felicidad”, oxitocina y dopamina, se multiplican en una parte del cerebro llamada cuerpo estriado ventral, lo que hace que los preadolescentes sean más sensibles a la atención y la admiración de los demás. A ello se le une la cercanía de otra zona denominada pálido ventral, clave en el sistema de motivación del comportamiento instintivo (Abrams, Zara, 2024). La acción de recibir un “Me gusta” en publicaciones proporciona a los usuarios una descarga de dopamina comparable a la experimentada con drogas y alcohol y potencia el comportamiento repetitivo (Nasti et al., 2021). Esta recompensa inmediata y visible refuerza el comportamiento de compartir contenido en busca de validación social, contribuyendo a la adicción. Los niños y adolescentes serían también más susceptibles experimentar síntomas de abstinencia al alejarse de las redes sociales. La participación en desafíos, el seguimiento de tendencias y la constante interacción social generan una conexión emocional profunda. La abstinencia de estas actividades puede desencadenar ansiedad, disforia, irritabilidad y comportamientos compulsivos, creando una dependencia emocional. La repetida



exposición a estímulos gratificantes podría llevar, además, a la neuroadaptación, tolerancia y a que los usuarios busquen incrementar su participación para experimentar la misma sensación placentera, creando un círculo vicioso de creciente dependencia (Montag & Diefenbach, 2018; Montag et al., 2019).

De otra parte, la inmadurez del cerebro de los niños y adolescentes influye en su capacidad para tomar decisiones informadas y de resistir las estrategias de diseño adictivo. La falta de desarrollo cognitivo completo hace que sean más susceptibles a las tácticas que explotan la gratificación instantánea y la búsqueda de validación (Murimi, 2023). Durante la adolescencia, el cerebro experimenta cambios significativos en áreas relacionadas con la toma de decisiones, la búsqueda de novedades y la regulación emocional. Además, el control de impulsos no está plenamente maduro, lo que dificulta detener comportamientos de beneficio inmediato, pero consecuencias negativas a largo plazo (American Psychological Association, 2024).

El uso intensivo de redes sociales en la adolescencia temprana se ha asociado con alteraciones en el desarrollo de la amígdala y el córtex prefrontal, ambos cruciales para el aprendizaje y control emocional y pudiera también generar una mayor sensibilidad a las recompensas y a las cancelaciones sociales (The U.S. Surgeon General's Advisory, 2023). Maza y colegas (2023) participants reported the frequency at which they checked Facebook, Instagram, and Snapchat. Neural responses to the Social Incentive Delay task when anticipating receiving social feedback, measured annually using fMRI for 3 years. Participants saw a cue that indicated whether the social feedback (adolescent faces with emotional expressions) sugieren que los comportamientos de revisión o chequeo constante de las redes sociales pueden estar asociados con cambios en la sensibilidad del cerebro a las recompensa y castigo sociales. Crone y Konijn (2018) concluyen que su mayor sensibilidad emocional y el desarrollo prolongado del procesamiento reflexivo y el control cognitivo pueden hacerlos específicamente reactivos a los medios que despiertan emociones y, por tanto, a las herramientas que se basan en la IA persuasiva. En la reciente revisión hecha por Ding y colegas (2023) se han concluido dos resultados principales. El primero es que los efectos perjudiciales de la DA en el cerebro de los niños y adolescentes hasta los 18 años son evidentes tanto en sus aspectos estructurales como funcionales. El segundo hallazgo involucra las regiones cerebrales específicas impactadas, que abarcan la corteza del cerebro (lóbulos frontal, parietal, temporal y occipital), el cerebelo y las estructuras subcorticales, como los ganglios basales (que contienen el estriado y el núcleo accumbens), el tálamo y el hipocampo. De todos ellos, la corteza prefrontal sería la región más vulnerable.



*Por su mayor vulnerabilidad psicológica: mayor sensibilidad al estrés social y mayor necesidad de validación y aprobación Social:*

Los niños y adolescentes están en un período de formación de su identidad y buscan constantemente la aprobación y validación de sus pares, siendo altamente sensibles hacia el rechazo social, de la misma manera que son altamente sensibles a la recompensa inmediata que ofrecen a través de “likes”, comentarios y seguidores... (Sherman et al., 2016), conformando sobre ello su idea de lo que socialmente es deseable. Se ha dicho que el cerebro adolescente (entre los 10-13 años y hasta los 25) tiene una hipersensibilidad al *feedback* social de sus pares, aumentando por ello su vulnerabilidad (Somerville, 2013, por todos).

La tendencia natural de los adolescentes a compararse con los otros se ve amplificada en entornos digitales. Las plataformas fomentan la comparación constante a través de imágenes, logros y estilo de vida, lo que puede llevar a sentimientos de insuficiencia y la necesidad de estar constantemente conectado (Irmer & Schmiedek, 2023) TikTok, and YouTube. El uso excesivo de redes sociales y la constante búsqueda de validación pueden contribuir a problemas de salud mental como ansiedad, depresión y baja autoestima, especialmente durante este período crucial de desarrollo personal y social (Twenge & Campbell, 2018). De otra parte, la naturaleza efímera del contenido y la generación constante de nuevos desafíos crean el temor a perderse experiencias significativas, que podrían estar disfrutando los demás, lo que es especialmente pronunciado en niños y adolescentes, que pueden sentirse excluidos si no están conectados constantemente (Przybylski et al., 2013; Oberst et al., 2017) but it has also been shown that these users can suffer from negative psychological consequences when using these sites excessively. We analyze the role of fear of missing out (FOMO) Además, investigaciones como la de Maza et al. (2023) participants reported the frequency at which they checked Facebook, Instagram, and Snapchat. Neural responses to the Social Incentive Delay task when anticipating receiving social feedback, measured annually using fMRI for 3 years. Participants saw a cue that indicated whether the social feedback (adolescent faces with emotional expressions sugieren que el comportamiento de revisión constante de redes sociales podría estar asociado con cambios en la sensibilidad del cerebro adolescente a las recompensas y castigos sociales. Es decir, que el acto de revisar las redes de forma repetitiva puede reforzar las respuestas cerebrales que hacen que los adolescentes sean más vulnerables a estímulos sociales como la aprobación o el rechazo. Esto puede generar un círculo vicioso, donde el adolescente siente la necesidad de mantenerse conectado para evitar “perderse” interacciones importantes o ser excluido socialmente.

De otra parte, se ha afirmado que los niños y adolescentes son también más receptivos y reactivos a la publicidad y a contenidos dirigidos y personali-



zados a través de inteligencia artificial, usando datos sobre la juventud y de los usuarios específicos. Los niños y adolescentes son especialmente susceptibles a las estrategias de marketing que se adaptan a sus preferencias y deseos, aumentando así su interacción con la plataforma. De la misma forma, son especialmente sensibles a contenidos polarizantes, altamente emocionales o de contenido lesivo. La razón es que están biológicamente predispuestos hacia la influencia de sus iguales y sensibles hacia ese contenido personalizado (American Psychological Association, 2024).

En resumen, la mayor vulnerabilidad de los adolescentes radica en las particularidades de su edad y de su desarrollo cerebral, concretadas en: su mayor sensibilidad a la valoración social, particularmente de sus pares; el desigual desarrollo de su cerebro, que les hace tener una mayor dificultad para la autoregulación emocional; y en la idoneidad de las redes sociales para activar y alterar los circuitos neuronales relacionados con el procesamiento de la recompensa (Costello et al., 2024).

#### 4. Potencial relevancia jurídico-penal de la persuasión algorítmica y del diseño adictivo con relación a los menores de edad

El objeto de este trabajo, tal y como se apuntaba al principio, es cuestionarse si el Derecho Penal habría de intervenir o no en el control de los riesgos derivados del diseño adictivo de las redes sociales y plataformas, a la vista de las conclusiones aportadas por la literatura científica especializada. Para valorar esta cuestión, desde una perspectiva de política-criminal, parece que el principio de intervención mínima obligaría a considerar legítimamente punibles únicamente aquellos riesgos a la salud pública de demostraran un potencial lesivo relevante y que, además, no hallaran una protección suficiente en otros sectores del ordenamiento jurídico, en particular, en el ámbito del Derecho Administrativo sancionador.

Desde esta perspectiva, ciertamente, no podría afirmarse que, a día de hoy, haya una posición dominante en la literatura científica en cuanto a la relación causal existente entre el deterioro de la salud mental y el uso abusivo de las redes sociales y plataformas similares, ni en adultos y ni siquiera entre adolescentes. The Lancet (2024) ha concluido que, hoy por hoy, los intentos de vincular de forma concluyente las tasas crecientes de enfermedad mental entre los jóvenes con el uso excesivo o compulsivo de las redes sociales está siendo complicado, al igual que hallar cambios cerebrales claros. Los resultados científicos hasta ahora, señalan, son mixtos, puesto que el uso de redes sociales también pueden tener efectos beneficiosos y porque en los problemas de salud mental de la ado-



lescencia podrían influir también otros factores que se sabe ya relevantes para la salud mental, como la desigualdad, el endurecimiento del mercado laboral, la preocupación por el cambio climático, la pobreza... (Tereshchenko, 2023). En el mismo sentido, en el estudio de Vaid et al., (2024). Además, Jaron Lenier (2018) ha precisado que la correlación entre el uso compulsivo de las redes sociales y los daños sobre la salud mental de los usuarios o su autonomía mental habría de considerarse realmente a nivel de grupo social, pero no podría establecerse de la misma forma, en términos de causalidad, con los usuarios individuales. Parece, pues, que la conclusión inicial, desde el punto de vista científico, sería que hace falta más investigación (Karim et al., 2020; Stangl et al., 2023; Irmer & Schmiedek, 2023 TikTok, and YouTube...), en condiciones que permitan sacar conclusiones fiables (Valkenburg et al., 2022) unified approach to investigate the effects of social media use on well-being. Using experience sampling methods among sizeable samples of respondents, our unified approach combines the strengths of nomothetic methods of analysis (e.g., mean comparisons, regression models; Brand, 2022...).

Al mismo tiempo, sin embargo, lo que no parece que pueda negarse tampoco es que existe un vasto cuerpo de literatura científica que sí indica claramente una correlación entre el deterioro de la salud mental de niños y adolescentes y el uso compulsivo de las redes sociales y plataformas similares, con un diseño adictivo y, muy en particular, en aquellas redes sociales basadas en imágenes y videos. Lo que no puede negarse tampoco, pues, es que muchos estudios están alertando insistentemente acerca de los riesgos graves que, para la salud mental de los más jóvenes, tendría este uso abusivo (Murimi, 2023). En el también muy reciente y muy documentado estudio de Costello y colegas (2024), con abundante bibliografía añadida a la que se recoge aquí, afirman sin rodeos la realidad de los daños que pueden provocar los algoritmos de las redes sociales en la salud mental de los niños y adolescentes, fundamentándose en el extenso cuerpo de literatura científica que avala con estudios experimentales y longitudinales esta asociación. También apuntan, con la misma contundencia, los importantes intereses económicos que mueven a las plataformas que usan este diseño adictivo y la necesidad de regular esta situación desde la perspectiva de la salud pública.

En mi opinión, en primer lugar, dado que los principales resultados lesivos y los riesgos derivados del diseño adictivo se están conectando con los menores de edad, en atención a la especial vulnerabilidad que presentan por su inmadurez neuropsicológica, el Derecho Penal debería intervenir, en su caso, limitando su intervención a la protección de dichas potenciales víctimas. De otra parte, considerando el objeto específico que estamos analizando, nos limitaremos a considerar en este planteamiento político-criminal la incidencia del diseño persuasivo algorítmico con la protección de la vida, a la salud individual



y pública y a la protección de los derechos de los consumidores menores de edad. Además, parece necesario revisar y cuestionarse seriamente la idea que se ha venido sostenido de la completa “*neutralidad*”, por definición, de las plataformas por los contenidos que difunden (Cfr. Silva Sánchez, 2023). Hasta ahora, se les ha venido considerando algo así como meras “librerías” respecto del contenido que contribuyen a difundir, sin considerar el modo real es que se produce esa difusión, mediante algoritmos persuasivos que permiten la manipulación del usuario y que son idóneos para generar adicción entre la juventud más vulnerable.

De otra parte, la valoración del nivel de riesgo permitido, cuando lo que está en peligro es la salud mental de la juventud, debería compararse cuidadosamente con el beneficio que el diseño adictivo de estas redes sociales y plataformas pudiera aportar a la sociedad. A mi juicio, este beneficio social, que permitiría aceptar un mayor umbral de riesgo, es nulo. Y por esta razón, la prioridad debe darse al principio de precaución y de responsabilidad. El único beneficio claro que concurre en el diseño adictivo de las redes sociales es para las propias empresas tecnológicas que las sustentan y para los anunciantes y otros clientes que se sirven de ellas. Es un negocio muy rentable para las plataformas, situándose en un volumen estimado de en EEUU de ingresos por publicidad dirigida hacia menores de edad de 0-17 de 11 mil millones de dólares en 2022 (Raffoul et al., 2023). No es de extrañar que las plataformas de redes sociales compitan ferozmente entre ellas por sus clientes, cuanto más jóvenes mejor, y que se imiten entre sí para incorporar características exitosas, creando un juego constante de “*Capture the Flag*” (Hurler, 2022), que retiene a los usuarios en un ciclo adictivo, dentro de un modelo de negocio de “*datos por uso de la aplicación*” (Montag et al., 2019).

Pues bien, en atención a todo lo anterior, la propuesta político-criminal que aquí se formula es que el Derecho Penal también debe intervenir para tutelar a los menores frente al diseño adictivo de las redes sociales y plataformas, basadas en voz, imagen y vídeo, sin perjuicio de la tutela administrativa que ya existe. Y ello, en atención a la relevancia de los riesgos generados y a la falta de valor social del diseño adictivo, que pudiera aconsejar modular en mayor medida el riesgo permitido o remitir la tutela al Derecho Administrativo. Esta propuesta normativa, más en concreto, habría de articularse sobre dos pilares fundamentales. El primero sería la preferencia de la racionalidad *pro-víctima*, frente a riesgos para los menores de edad, inciertos y con potencial sistémico. A partir de ahí, construir la protección penal sobre la base del principio de precaución y de responsabilidad, priorizando, por tanto, el “control del riesgo” ante la previsibilidad de unos daños y riesgos que la literatura conecta al diseño adictivo de las redes sociales y de las plataformas (peligrosidad *ex ante* de la conducta). El segundo pilar, considerar estos diseños adictivos fuera del riesgo permitido con



relación a los menores, sirviéndonos del *paraguas* que proporciona ya el art. 5 del Reglamento IA de la UE y centrándonos muy particularmente, en la responsabilidad de las plataformas que usan tales herramientas de IA para conseguir la atención y la constante y compulsiva conexión del usuario.

- 4.1. Potencial responsabilidad por los daños generados por la exposición constante a contenidos lesivos, motivada o favorecida por el diseño algorítmico de redes sociales, plataformas de comunicación o, en general, espacios o servicios virtuales.

Conforme al Código Penal vigente, los daños que puedan causarse a un menor en su salud mental, como consecuencia de su adicción a las redes sociales, o los que, como resultado de ello, pudiera autoinfligirse a sí mismo, son atípicos. La tipicidad de esta clase de comportamientos, tratados como delitos de resultado, plantearía enormes dificultades con relación la prueba de la relación de causalidad entre el diseño adictivo de la plataforma (aún más sin acceso al algoritmo) y el daño provocado a un sujeto concreto. Podría oponerse fácilmente que, en puridad, el daño se conectaría más directamente con el contenido peligroso que con el propio diseño adictivo, que sólo podría suponer el incremento de un riesgo principal creado por la aportación de otro o, más propiamente, por los contenidos de otros muchos, no concertados. El algoritmo sólo habría proporcionado o sugerido los contenidos lesivos que otros habrían incorporado a la red. Consecuentemente, también sería extremadamente difícil establecer la responsabilidad criminal, incluso a título de imprudencia, conforme a los parámetros tradicionales y máxime, cuando concurren diferencias individuales muy significativas en la potencial afectación de los adolescentes concretos.

De otra parte, aun cuando lo anterior pudiera sortearse, faltaría la tipificación expresa de la responsabilidad penal de las personas jurídicas en el delito de auxilio o inducción al suicidio y en los delitos contra la salud. Este problema se ha planteado en varias demandas en EEUU, hasta ahora sin éxito, puesto que los tribunales han adoptado una postura que equipara a las plataformas digitales con proveedores de servicios, atribuyendo una limitada responsabilidad por los contenidos que circulan en sus redes. [Así en *Twitter, Inc. v. Taamneh*, 143 S.Ct. 1206, 1209 (2023) y *Gonzalez v. Google LLC*, 143 S. Ct. 1191 (2023)]. En la misma línea, se está todavía sustanciando otra demanda colectiva en la que se reclama específicamente a las plataformas por los daños producidos a sus usuarios, derivados de su diseño adictivo. Así, en el *Caso 2:23-cv-00032 - Seattle Public Schools vs. Facebook, Instagram de Meta; Snapchat, Tik Tok y YouTube* se acusa a estas empresas de afectar la salud mental de los jóvenes. La analogía con las “nuevas tabaqueras” sugiere que estas plataformas, de manera similar a las compañías tabacaleras, habrían contribuido a problemas de salud, en lugar de prevenirlos.



#### 4.2. Potencial responsabilidad como delitos de peligro para la salud de las personas o salud pública

En el ámbito de los delitos contra las personas, por la Disposición Final 6<sup>a</sup> de la LO 8/2021, de 4 de junio, se introdujeron en el Código Penal, en lo que aquí nos interesa, los artículos 143bis), con relación a la inducción o favorecimiento del suicidio en menores o personas con discapacidad y el 156ter, referido a las autolesiones. Así, por ejemplo, el primero de ellos, castiga con una pena de prisión, como forma de inducción o favorecimiento del suicidio de un menor o persona vulnerable por razón de su discapacidad: “la distribución o difusión pública a través de internet, del teléfono o de cualquier otra tecnología de la información o de la comunicación de contenidos específicamente destinados a promover, fomentar o incitar al suicidio de personas menores de edad”.

La tipificación de estos comportamientos es insuficiente, puesto que en ambos casos se atiende exclusivamente al contenido lesivo o inapropiado que se difunde (que debe ser, además, específico) y no también al diseño algorítmico de la plataforma o medio digital que contribuye a esa difusión, a través del “contenido sugerido”. De otra parte, parece circunscribir la tipicidad a los comportamientos dirigidos específicamente a favorecer tales comportamientos autolesivos, lo que parece requerir dolo directo y dejar fuera (o dificultar en gran medida) “*el efecto razonablemente probable*”, que estos pudieran provocar en estos colectivos vulnerables. Dejarían fuera, por tanto, también el dolo eventual.

Además, con esta tipificación pareciera que se acoge la idea de que la plataforma o servicio a través del que se difunde ese contenido está realizando una conducta neutral, dada la pena de prisión prevista y la falta de referencia a la responsabilidad de las personas jurídicas. Es decir, se considera que la plataforma, por ejemplo TikTok, actuaría como mero servicio intermediario no como creador de contenidos, presumiendo su buena fe o, cuando menos, su falta de responsabilidad por la difusión de tales contenidos, amparado por una conducta profesional o empresarial dentro del riesgo permitido. Deliberadamente se ha ignorado, no sólo que quien haya creado el contenido lesivo pueda ser también una persona jurídica, sino que quien lo difunde o distribuye públicamente sólo puede hacerlo, porque, tras la persona física, existe una persona jurídica que le presta su infraestructura digital, que se lucra con la difusión de dichos contenidos a través de diversas fórmulas, que puede llegar a tener constancia (o deliberadamente desentenderse) del carácter lesivo de tales contenidos que está difundiendo y sugiriendo repetida y personalmente.

Dentro de los delitos contra salud pública, se ha introducido también el art. 361 bis, que, con similar redacción, castiga a quien difunda o distribuya contenidos específicamente destinados promover los trastornos o riesgos alimenti-



cios: “consumo de productos, preparados o sustancias o la utilización de técnicas de ingestión o eliminación de productos alimenticios cuyo uso sea susceptible de generar riesgo para la salud de las personas”. Se castiga con pena de multa de 6 a 12 meses o prisión de 1 a 3 años, salvo que sea aplicable el 362<sup>quater</sup> Cp. Si fuera una persona jurídica, es de aplicación el art. 366 Cp, cuya previsión penológica sería complemente inadecuada para estos supuestos, que requeriría, por ejemplo, atender al volumen de negocio o de usuarios que una determinada plataforma digital pudiera tener en España para cuantificar la multa. Parece, de nuevo, que ello responde a considerar que las personas jurídicas que pudieran ser responsables por este delito son aquellas que crean esos contenidos o puedan lucrarse con los mismos o con la venta de sustancias, preparados o productos... y no por aquellas plataformas que se limitan a difundir a través de ellas dichos contenidos o sugerirlos de forma personalizada. La misma premisa político-criminal subyacente en los art. 143 bis y 156ter del Código Penal.

En mi opinión, la descripción típica habría de ampliarse para incorporar la eventual responsabilidad penal de las personas jurídicas que con sus diseños o productos contribuyan también a la difusión y distribución de tales contenidos entre usuarios indeterminados o por difusión personalizada, cuando parezca razonable esperar, que pueden ser accesibles a menores de edad y no se hayan tomado medidas efectivas para controlar este riesgo. Lo determinante será, para que tal tipificación sea algo más que simbólica y pueda probarse (Costello et al., 2024), que se haga exigible auditorías de riesgos independientes y públicas a estas *gatekeepers*, conforme a lo que establece el Reglamento (UE) 2022/1925 del Parlamento Europeo y del Consejo, de 14 de septiembre de 2022, relativo a los mercados digitales (Digital Markets Act - DMA), entre otras responsabilidades y obligaciones. Por ejemplo, el “sistema de recompensas variables intermitentes” ha motivado una investigación iniciada por la Comisión de la Unión Europea a TikTok por la creación de “*TikTok Lite Rewards*”, que premiaba a los usuarios por tiempo y contenido consumido en la plataforma. Dicha versión de la app fue retirada, subsistiendo todavía otras por la falta de transparencia en la gestión de sus anuncios, por la falta de eficacia de control de acceso y por no poner medidas en la gestión de riesgos de su diseño adictivo con respecto a los menores (Europa Press, 2024).

Pero, además, como delito contra la salud pública, habría también de tipificarse expresamente el uso de diseños adictivos en plataformas, redes sociales y apps... que estén dirigidas específicamente a usuarios menores de edad o usuarios indeterminados, cuando, de nuevo, parezca razonable esperar que puedan ser accesibles a menores de edad y no se hayan tomado medidas efectivas para controlar este riesgo. Debiera prohibirse la utilización de IA susceptible de causar adicción en la población y más en concreto, entre los menores de edad. El diseño adictivo de las plataformas, particularmente cuando se utilizan



recompensas variables intermitentes, no tiene utilidad social alguna; sólo tiene una utilidad económica para las empresas que los utilizan o se sirven de ello. Por tanto, no puede entenderse que deba ponerse en riesgo la salud mental de los más jóvenes, menos aún, cuando hay multitud de estudios científicos que están avisando de ello, para salvaguardar un entorno algorítmico que no reporta ninguna utilidad social.

Creo además, como se ha señalado, que esta conclusión podría fundamentarse en lo dispuesto en el art. 5.1 del Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024 (en adelante, Reglamento IA). En dicho precepto se consideran prohibidas determinadas prácticas de manipulación y de explotación posibilitadas por la IA y, entre ellas, en su apartado b) proscribire: *“b) la introducción en el mercado, la puesta en servicio o la utilización de un sistema de IA que explote alguna de las vulnerabilidades de una persona física o un determinado colectivo de personas derivadas de su edad o discapacidad, o de una situación social o económica específica, con la finalidad o el efecto de alterar de manera sustancial el comportamiento de dicha persona o de una persona que pertenezca a dicho colectivo de un modo que provoque, o sea razonablemente probable que provoque, perjuicios considerables a esa persona o a otra.”*

Conforme a los datos y argumentos que se han venido exponiendo, el diseño adictivo de las redes sociales explota deliberadamente las vulnerabilidades del ser humano y, muy en particular, de los niños y adolescentes, siendo perfectamente idóneas, como señala la literatura científica, para alterar sustancialmente el comportamiento de éstos, hasta generarles, siquiera en un porcentaje indeterminado de ellos, adicción y otros daños a su salud mental por sobreexposición a los contenidos lesivos distribuidos por ellas. Y como se explica en el Reglamento de la IA en su considerando 29, “no puede presuponerse que existe la intención de alterar el comportamiento si la alteración es el resultado de factores externos al sistema de IA que escapan al control del proveedor o del responsable del despliegue” (p.e. otras vulnerabilidades específicas de un determinado sujeto), pero tampoco es necesario que el proveedor o el responsable del despliegue tengan la intención de causar un perjuicio considerable, siempre que dicho perjuicio se derive de las prácticas de manipulación o explotación que posibilita la IA.” y que, por tanto, pudieran ser previsibles, conforme a la auditoría de riesgos que vienen obligadas a realizar. Y el peligro de generar adicción entre los más jóvenes y los riesgos que de ello pueden derivarse, siquiera para un sector de esta población, creo que hoy por hoy, no puede simplemente desconocerse.



#### 4.3. Potencial responsabilidad por el producto en entornos o servicios digitales, como delitos contra los consumidores

Esta exigencia de responsabilidad por el producto y de protección a los consumidores menores de edad es el fundamento central de otra de las demandas colectivas que se están tramitando en EEUU contra las grandes empresas tecnológicas. Caso 4:22-md-03047-YGR *-In re: Social Media Adolescent Addiction/ Personal Injury Products Liability Litigation MDL*. Se trata de una demanda en la que se reclama la responsabilidad de gigantes tecnológicos como Google, Meta, ByteDance y Snap por diseñar sus productos conscientemente para enganchar a los usuarios, a pesar de afirmar que sus redes sociales son seguras para los menores o, cuando menos, habiendo hecho un diseño defectuoso por imprudencia o habiendo fallado a la hora de advertir de los potenciales daños que pudieran derivarse para estos consumidores. La base de la demanda es la presunta manipulación de las plataformas para atraer a un público más joven, a expensas de su bienestar, aprovechando el sistema de recompensas químicas en el cerebro de los usuarios, especialmente los jóvenes, induciendo compromiso adictivo, uso compulsivo y daños mentales y físicos. Se destaca, además, el “efecto máquina tragaperras” como una estrategia empleada por Meta (anteriormente Facebook) para ofrecer contenido algorítmicamente a los usuarios jóvenes, manipulando la liberación de dopamina y induciéndolos injusta o deslealmente a utilizar repetidamente sus productos.

En el ámbito socioeconómico de la protección de los consumidores no hay tampoco ninguna previsión específica ni adecuada en el Código Penal vigente. Así, el art. 282 Cp., que regula la prohibición de publicidad falsa, no contiene previsión alguna respecto del uso de técnicas subliminales y, además, tampoco se hace referencia a la responsabilidad de las plataformas por su propio producto, que estaría diseñado mediante herramientas de IA persuasiva, con algoritmos dirigidos a fomentar el uso intensivo, compulsivo y adictivo de su producto, también entre los consumidores menores de edad, explotando o, cuando menos, ignorando, su especial vulnerabilidad y necesidades de protección.

Utilizar servicios digitales con diseños adictivos mediados con IA, como instrumento captador de atención y compromiso del usuario menor de edad para, a partir de él, introducir mensajes manipulativos o publicitarios, son prácticas abusivas para el consumidor, como prácticas de naturaleza subliminales o, cuando menos, engañosas. Además, no podría considerarse una práctica comercial común y legítima, cuando dicha publicidad sea dirigida y/o fácilmente accesible, para los menores de edad, sin atender a las limitaciones legales en este ámbito. En esta línea se pronuncian Esposito y Ferreira (2024), considerando que el diseño adictivo debe ser considerado ilegal y una práctica comercial ilícita en el marco de la Unión Europea. De la misma forma, debería incorporarse



una tipificación penal expresa con respecto a publicidad subliminal a través de realidad virtual (González-Tapia, 2023).

## 5. Bibliografía

- Abrams, Zara. (2024). *Why young brains are especially vulnerable to social media*. <https://www.apa.org/news/apa/2022/social-media-children-teens>
- Alter, A. (2017). *Irresistible: The rise of addictive technology and the business of keeping us hooked* (p. 354). Penguin Press.
- American Psychological Association. (2024). *Potential Risks of Content, Features, and Functions: A closer look at the science behind how social media affects youth*. <https://www.apa.org/topics/social-media-internet/psychological-science-behind-youth-social-media.pdf>
- Anderson, M., Faverio, M., & Gottfried, J. (2023). *Teens, Social Media and Technology 2023*.
- Andreassen, C. S., Pallesen, S., & Griffiths, M. D. (2017). The relationship between addictive use of social media, narcissism, and self-esteem: Findings from a large national survey. *Addictive Behaviors*, 64, 287-293. <https://doi.org/10.1016/j.addbeh.2016.03.006>
- Beyens, I., Pouwels, J. L., van Driel, I. I., Keijsers, L., & Valkenburg, P. M. (2020). The effect of social media on well-being differs from adolescent to adolescent. *Scientific Reports*, 10(1), 10763. <https://doi.org/10.1038/s41598-020-67727-7>
- Billieux, J., Schimmenti, A., Khazaal, Y., Maurage, P., & Heeren, A. (2015). Are we overpathologizing everyday life? A tenable blueprint for behavioral addiction research. *Journal of Behavioral Addictions*, 4(3), 119-123. <https://doi.org/10.1556/2006.4.2015.009>
- Bloomberg.com. (2022). Parents Suing Social Media Companies Blame Sites for Getting Kids Addicted. *Bloomberg.Com*. <https://www.bloomberg.com/news/articles/2022-09-22/social-media-addiction-brings-lawsuits-against-youtube-facebook>
- Brand, M. (2022). Can internet use become addictive? *Science*, 376(6595), 798-799. <https://doi.org/10.1126/science.abn4189>
- Carr, C. T., & Rosaen, S. F. (2024). We're Going Striking!: Associations Between the Gamification of Mediated Communication and Relational



- Closeness. *Communication Reports*, 37(3), 185-199. <https://doi.org/10.1080/08934215.2024.2334743>
- CCDH. (2024). *TikTok removes views from hashtags*. Center for Countering Digital Hate | CCDH. <https://counterhate.com/research/tiktok-removes-views-from-hashtags/>
- Channel 4 News (Director). (2018, junio 15). *Jaron Lanier interview on how social media ruins your life* [Video recording]. [https://www.youtube.com/watch?v=kc\\_Jq42Og7Q](https://www.youtube.com/watch?v=kc_Jq42Og7Q)
- Chen, W., & Lee, K.-H. (2013). Sharing, Liking, Commenting, and Distressed? The Pathway Between Facebook Interaction and Psychological Distress. *Cyberpsychology, Behavior, and Social Networking*, 16(10), 728-734. <https://doi.org/10.1089/cyber.2012.0272>
- Chen, Y., Li, M., Guo, F., & Wang, X. (2023). The effect of short-form video addiction on users' attention. *Behaviour & Information Technology*, 42(16), 2893-2910. <https://doi.org/10.1080/0144929X.2022.2151512>
- Clark, L., & Zack, M. (2023). Engineered highs: Reward variability and frequency as potential prerequisites of behavioural addiction. *Addictive Behaviors*, 140, 107626. <https://doi.org/10.1016/j.addbeh.2023.107626>
- Costello, N., Sutton, R., Jones, M., Almassian, M., Raffoul, A., Ojumu, O., Salvia, M., Santoso, M., Kavanaugh, J. R., & Austin, S. B. (2024). ALGORITHMS, ADDICTION, AND ADOLESCENT MENTAL HEALTH: An Interdisciplinary Study to Inform State-level Policy Action to Protect Youth from the Dangers of Social Media. *American Journal of Law & Medicine*, 49(2-3), 135-172. <https://doi.org/10.1017/amj.2023.25>
- Crone, E. A., & Konijn, E. A. (2018). Media use and brain development during adolescence. *Nature Communications*, 9(1), 588. <https://doi.org/10.1038/s41467-018-03126-x>
- Delic, A., Delfabbro, P., & King, D. L. (2023). Understanding the Risks and Individual Difference Factors Associated with Engagement in Monetised (Play-to-Earn) Gaming. *International Journal of Mental Health and Addiction*. <https://doi.org/10.1007/s11469-023-01105-y>
- Ding, K., Shen, Y., Liu, Q., & Li, H. (2023). The Effects of Digital Addiction on Brain Function and Structure of Children and Adolescents: A Scoping Review. *Healthcare*, 12(1), 15. <https://doi.org/10.3390/healthcare12010015>



- Dino Ambrosi (Director). (2023, marzo 6). *The Battle for Your Time: Exposing the Costs of Social Media* | Dino Ambrosi | TEDxLagunaBlancaSchool [Video recording]. <https://www.youtube.com/watch?v=4TMPXK9tw5U>
- Dyer, C. (2022). Social media content contributed to teenager's death "in more than a minimal way," says coroner. *BMJ*, 379, o2374. <https://doi.org/10.1136/bmj.o2374>
- Elhai, J. D., Yang, H., & Montag, C. (2019). Cognitive- and Emotion-Related Dysfunctional Coping Processes: Transdiagnostic Mechanisms Explaining Depression and Anxiety's Relations with Problematic Smartphone Use. *Current Addiction Reports*, 6(4), 410-417. <https://doi.org/10.1007/s40429-019-00260-4>
- Esposito, F., & Ferreira, T. M. C. (2024). Addictive Design as an Unfair Commercial Practice: The Case of Hyper-Engaging Dark Patterns. *European Journal of Risk Regulation*, 1-18. <https://doi.org/10.1017/err.2024.8>
- Europa Press. (2024, febrero 19). La Comisión Europea investigará a TikTok por falta de transparencia y por no proteger a los menores de su diseño adictivo. *El País*. <https://elpais.com/sociedad/2024-02-19/la-comision-europea-investigara-a-tiktok-por-falta-de-transparencia-y-por-no-proteger-a-los-menores-de-su-diseno-adictivo.html>
- Farchakh, Y., Haddad, C., Sacre, H., Obeid, S., Salameh, P., & Hallit, S. (2020). Video gaming addiction and its association with memory, attention and learning skills in Lebanese children. *Child and Adolescent Psychiatry and Mental Health*, 14(1), 46. <https://doi.org/10.1186/s13034-020-00353-3>
- Firth, J., Torous, J., Stubbs, B., Firth, J. A., Steiner, G. Z., Smith, L., Alvarez-Jimenez, M., Gleeson, J., Vancampfort, D., Armitage, C. J., & Sarris, J. (2019). The "online brain": How the Internet may be changing our cognition. *World Psychiatry*, 18(2), 119-129. <https://doi.org/10.1002/wps.20617>
- Fuhrmann, D., Knoll, L. J., & Blakemore, S.-J. (2015). Adolescence as a Sensitive Period of Brain Development. *Trends in Cognitive Sciences*, 19(10), 558-566. <https://doi.org/10.1016/j.tics.2015.07.008>
- GEM UNESCO. (2024). *Global education monitoring report 2024, gender report: Technology on her terms*. GEM Report UNESCO. <https://doi.org/10.54676/WVCF2762>
- Gingras, M.-P., Brendgen, M., Beauchamp, M. H., Séguin, J. R., Tremblay, R. E., Côté, S. M., & Herba, C. M. (2024). Adolescents and Social Me-



- dia: Longitudinal Links Between Motivations for Using Social Media and Subsequent Internalizing Symptoms. *Journal of Youth and Adolescence*. <https://doi.org/10.1007/s10964-024-02097-1>
- González-Tapia, M. I. (2023). Virtual emotions and Criminal Law. *Frontiers in Psychology*, 14, 1260425. <https://doi.org/10.3389/fpsyg.2023.1260425>
- Griffiths, M. D. (2018). *Adolescent social networking: How do social media operators facilitate habitual use?*
- Huang, S., Lai, X., Ke, L., Li, Y., Wang, H., Zhao, X., Dai, X., & Wang, Y. (2024). AI Technology panic—is AI Dependence Bad for Mental Health? A Cross-Lagged Panel Model and the Mediating Roles of Motivations for AI Use Among Adolescents. *Psychology Research and Behavior Management*, 17, 1087-1102. <https://doi.org/10.2147/PRBM.S440889>
- Hurler, K. (2022, agosto 16). *For Sites Like Instagram and Twitter, Imitation Is the Only Form of Flattery*. Gizmodo. <https://gizmodo.com/instagram-tiktok-snapchat-facebook-meta-1849395419>
- Irmer, A., & Schmiedek, F. (2023). Associations between youth's daily social media use and well-being are mediated by upward comparisons. *Communications Psychology*, 1(1), 1-13. <https://doi.org/10.1038/s44271-023-00013-0>
- Karim, F., Oyewande, A., Abdalla, L. F., Chaudhry Ehsanullah, R., & Khan, S. (2020). Social Media Use and Its Connection to Mental Health: A Systematic Review. *Cureus*. <https://doi.org/10.7759/cureus.8627>
- Keipi, T., Näsi, M., Oksanen, A., & Räsänen, P. (2016). *Online Hate and Harmful Content: Cross-national perspectives* (1.<sup>a</sup> ed.). Routledge. <https://doi.org/10.4324/9781315628370>
- King, D. L., & Delfabbro, P. H. (2020). Video game addiction. En *Adolescent Addiction* (pp. 185-213). Elsevier. <https://doi.org/10.1016/B978-0-12-818626-8.00007-4>
- Kuss, D. J., & Griffiths, M. D. (2017). Social Networking Sites and Addiction: Ten Lessons Learned. *International Journal of Environmental Research and Public Health*, 14(3), Article 3. <https://doi.org/10.3390/ijerph14030311>
- Kuss, D. J., Kristensen, A. M., & Lopez-Fernandez, O. (2021). Internet addictions outside of Europe: A systematic literature review. *Computers in Human Behavior*, 115, 106621. <https://doi.org/10.1016/j.chb.2020.106621>
- Lee, D., Namkoong, K., Lee, J., & Jung, Y. (2021). Dorsal striatal functional connectivity changes in Internet gaming disorder: A longitudinal magnetic



- resonance imaging study. *Addiction Biology*, 26(1), e12868. <https://doi.org/10.1111/adb.12868>
- Lonergan, A. R., Bussey, K., Fardouly, J., Griffiths, S., Murray, S. B., Hay, P., Mond, J., Trompeter, N., & Mitchison, D. (2020). Protect me from my selfie: Examining the association between photo-based social media behaviors and self-reported eating disorders in adolescence. *International Journal of Eating Disorders*, 53(5), 755-766. <https://doi.org/10.1002/eat.23256>
- Lopez-Fernandez, O., Romo, L., Rousseau, A., Lelonek-Kuleta, B., Chwaszcz, J., Männikkö, N., Gässler, A.-K., Demetrovics, Z., Achab, S., Kuss, D. J., & Griffiths, M. D. (2024). Problematic Internet use among adults: A longitudinal European study. *Adicciones*. <https://doi.org/10.20882/adicciones.1948>
- Lozano-Muñoz, N., Borrallo-Riego, Á., Guerra-Martín, M. D., Lozano-Muñoz, N., Borrallo-Riego, Á., & Guerra-Martín, M. D. (2022). Influencia de las redes sociales sobre la anorexia y la bulimia en las adolescentes: Una revisión sistemática. *Anales del Sistema Sanitario de Navarra*, 45(2). <https://doi.org/10.23938/asn.1009>
- Matz, S. C., Kosinski, M., Nave, G., & Stillwell, D. J. (2017). Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the National Academy of Sciences*, 114(48), 12714-12719. <https://doi.org/10.1073/pnas.1710966114>
- Matz, S. C., Teeny, J. D., Vaid, S. S., Peters, H., Harari, G. M., & Cerf, M. (2024). The potential of generative AI for personalized persuasion at scale. *Scientific Reports*, 14(1), 4692. <https://doi.org/10.1038/s41598-024-53755-0>
- Maza, M. T., Fox, K. A., Kwon, S.-J., Flannery, J. E., Lindquist, K. A., Prinstein, M. J., & Telzer, E. H. (2023). Association of Habitual Checking Behaviors on Social Media With Longitudinal Functional Brain Development. *JAMA Pediatrics*, 177(2), 160-167. <https://doi.org/10.1001/jamapediatrics.2022.4924>
- Montag, C., & Diefenbach, S. (2018). Towards Homo Digitalis: Important Research Issues for Psychology and the Neurosciences at the Dawn of the Internet of Things and the Digital Society. *Sustainability*, 10(2), Article 2. <https://doi.org/10.3390/su10020415>
- Montag, C., Lachmann, B., Herrlich, M., & Zweig, K. (2019). Addictive Features of Social Media/Messenger Platforms and Freemium Games against the Background of Psychological and Economic Theories. *International*



*Journal of Environmental Research and Public Health*, 16(14), 2612. <https://doi.org/10.3390/ijerph16142612>

- Morgans, J. (2017, mayo 18). The Secret Ways Social Media Is Built for Addiction. *VICE*. <https://www.vice.com/en/article/the-secret-ways-social-media-is-built-for-addiction/>
- Murimi, R. (2023). What Makes AI Addictive? The Role of Discounting, Risk Aversion and Self-regulation. En C.-H. Hsu, M. Xu, H. Cao, H. Baghban, & A. B. M. Shawkat Ali (Eds.), *Big Data Intelligence and Computing* (pp. 451-462). Springer Nature. [https://doi.org/10.1007/978-981-99-2233-8\\_32](https://doi.org/10.1007/978-981-99-2233-8_32)
- Nasti, L., Michienzi, A., & Guidi, B. (2021). Discovering the Impact of Notifications on Social Network Addiction. En J. Bowles, G. Broccia, & M. Nanni (Eds.), *From Data to Models and Back* (Vol. 12611, pp. 72-86). Springer International Publishing. [https://doi.org/10.1007/978-3-030-70650-0\\_5](https://doi.org/10.1007/978-3-030-70650-0_5)
- Nesi, J., & Prinstein, M. J. (2015). Using Social Media for Social Comparison and Feedback-Seeking: Gender and Popularity Moderate Associations with Depressive Symptoms. *Journal of Abnormal Child Psychology*, 43(8), 1427-1438. <https://doi.org/10.1007/s10802-015-0020-0>
- Nogueira-López, A., Rial-Boubeta, A., Guadix-García, I., Villanueva-Blasco, V. J., & Billieux, J. (2023). Prevalence of problematic Internet use and problematic gaming in Spanish adolescents. *Psychiatry Research*, 326, 115317. <https://doi.org/10.1016/j.psychres.2023.115317>
- Oberst, U., Wegmann, E., Stodt, B., Brand, M., & Chamarro, A. (2017). Negative consequences from heavy social networking in adolescents: The mediating role of fear of missing out. *Journal of Adolescence*, 55(1), 51-60. <https://doi.org/10.1016/j.adolescence.2016.12.008>
- O'Brien, O., Sumich, A., Baguley, T., & Kuss, D. J. (2023). A partial correlation network indicates links between wellbeing, loneliness, FOMO and problematic internet use in university students. *Behaviour & Information Technology*, 42(16), 2717-2734. <https://doi.org/10.1080/0144929X.2022.2142845>
- OEDA, C. (2023). *Informe sobre Adicciones comportamentales y otros trastornos adictivos 2023*. Observatorio Español de las Drogas y las Adicciones Delegación del Gobierno para el Plan Nacional sobre Drogas.
- Oksanen, A., Celuch, M., Oksa, R., & Savolainen, I. (2024). Online communities come with real-world consequences for individuals and societies.



- Communications Psychology*, 2(1), 1-9. <https://doi.org/10.1038/s44271-024-00112-6>
- ONTSI. (2023). *Impacto del aumento del uso de Internet y las redes sociales en la salud mental de jóvenes y adolescentes*. <https://www.ontsi.es/sites/ontsi/files/2023-10/policybriefredesocialesaludmentaljovenesyadolescentes.pdf>
- ONTSI. (2024). *El uso de las tecnologías por menores en España*. Observatorio Nacional de Tecnología y Sociedad. <https://doi.org/10.30923/23024007X>
- Przybylski, A. K., Murayama, K., DeHaan, C. R., & Gladwell, V. (2013). Motivational, emotional, and behavioral correlates of fear of missing out. *Computers in Human Behavior*, 29(4), 1841-1848. <https://doi.org/10.1016/j.chb.2013.02.014>
- Raffoul, A., Ward, Z. J., Santoso, M., Kavanaugh, J. R., & Austin, S. B. (2023). Social media platforms generate billions of dollars in revenue from U.S. youth: Findings from a simulated revenue model. *PLOS ONE*, 18(12), e0295337. <https://doi.org/10.1371/journal.pone.0295337>
- Rideout, V. (2021). *The Common Sense Census: Media Use by Tweens and Teens in America, A Common Sense Media Research Study, United States, 2015: Version 1 (Versión v1) [Dataset]*. ICPSR - Interuniversity Consortium for Political and Social Research. <https://doi.org/10.3886/ICPSR38018.V1>
- Romer, D. (2010). Adolescent risk taking, impulsivity, and brain development: Implications for prevention. *Developmental Psychobiology*, 52(3). <https://doi.org/10.1002/dev.20442>
- Ryding, F. C., Harkin, L. J., & Kuss, D. J. (2024). Instagram engagement and well-being: The mediating role of appearance anxiety. *Behaviour & Information Technology*, 1-17. <https://doi.org/10.1080/0144929X.2024.2323078>
- Savolainen, I., & Oksanen, A. (2024). Keeping you connected or keeping you addicted? Weekly use of social media platforms is associated with hazardous alcohol use and problem gambling among adults. *Alcohol and Alcoholism*, 59(3), agae024. <https://doi.org/10.1093/alcalc/ague024>
- Şenol, Y., Şenol, F. B., & Can Yaşar, M. (2024). Digital game addiction of pre-school children in the Covid-19 pandemic: Social emotional development and parental guidance. *Current Psychology*, 43(1), 839-847. <https://doi.org/10.1007/s12144-023-04323-8>



- Shanmugasundaram, M., & Tamilarasu, A. (2023). The impact of digital technology, social media, and artificial intelligence on cognitive functions: A review. *Frontiers in Cognition*, 2. <https://doi.org/10.3389/fcogn.2023.1203077>
- Shannon, H., Bush, K., Villeneuve, P.J., Hellemans, K. G., & Guimond, S. (2022). Problematic Social Media Use in Adolescents and Young Adults: Systematic Review and Meta-analysis. *JMIR Mental Health*, 9(4), e33450. <https://doi.org/10.2196/33450>
- Sherman, L. E., Payton, A. A., Hernandez, L. M., Greenfield, P. M., & Dapretto, M. (2016). The Power of the *Like* in Adolescence: Effects of Peer Influence on Neural and Behavioral Responses to Social Media. *Psychological Science*, 27(7), 1027-1035. <https://doi.org/10.1177/0956797616645673>
- Silva Sánchez, J. M. (2023). ¿Hasta qué punto son conductas neutras los servicios de Google, Facebook o Twitter? – InDret. *Indret*, 3, 3.
- Singer, N. (2015, diciembre 5). Can't Put Down Your Device? That's by Design. *The New York Times*. <https://www.nytimes.com/2015/12/06/technology/personaltech/cant-put-down-your-device-thats-by-design.html>
- Somerville, L. H. (2013). Special issue on the teenage brain: Sensitivity to social evaluation. *Current Directions in Psychological Science*, 22(2), 121. <https://doi.org/10.1177/0963721413476512>
- Soriano-Sánchez, J. (2022). Factores psicológicos y consecuencias del Síndrome Fear of Missing Out: Una Revisión Sistemática. *Revista de Psicología y Educación - Journal of Psychology and Education*, 17, 69. <https://doi.org/10.23923/rpye2022.01.217>
- Stangl, F. J., Riedl, R., Kiemeswenger, R., & Montag, C. (2023). Negative psychological and physiological effects of social networking site use: The example of Facebook. *Frontiers in Psychology*, 14, 1141663. <https://doi.org/10.3389/fpsyg.2023.1141663>
- Sumner, S. A., Ferguson, B., Bason, B., Dink, J., Yard, E., Hertz, M., Hilker, B., Holland, K., Mercado-Crespo, M., Tang, S., & Jones, C. M. (2021). Association of Online Risk Factors With Subsequent Youth Suicide-Related Behaviors in the US. *JAMA Network Open*, 4(9), e2125860. <https://doi.org/10.1001/jamanetworkopen.2021.25860>
- Tereshchenko, S. Y. (2023). Neurobiological risk factors for problematic social media use as a specific form of Internet addiction: A narrative review. *World Journal of Psychiatry*, 13(5), 160-173. <https://doi.org/10.5498/wjp.v13.i5.160>



- The Lancet. (2024). Unhealthy influencers? Social media and youth mental health. *The Lancet*, 404(10461), 1375. [https://doi.org/10.1016/S0140-6736\(24\)02244-X](https://doi.org/10.1016/S0140-6736(24)02244-X)
- The U.S. Surgeon General's Advisory. (2023). *Social Media and Youth Mental Health*. U.S. Department of Health and Human Services. Office of the Surgeon General. <https://www.hhs.gov/surgeongeneral/priorities/youth-mental-health/social-media/index.html>
- Towner, E., Chierchia, G., & Blakemore, S.-J. (2023). Sensitivity and specificity in affective and social learning in adolescence. *Trends in Cognitive Sciences*, 27(7), 642-655. <https://doi.org/10.1016/j.tics.2023.04.002>
- Turel, O., He, Q., Xue, G., Xiao, L., & Bechara, A. (2014a). Examination of Neural Systems Sub-Serving Facebook "Addiction". *Psychological Reports*, 115(3), 675-695. <https://doi.org/10.2466/18.PR.0.115c31z8>
- Turel, O., He, Q., Xue, G., Xiao, L., & Bechara, A. (2014b). Examination of Neural Systems Sub-Serving Facebook "Addiction". *Psychological Reports*, 115(3), 675-695. <https://doi.org/10.2466/18.PR.0.115c31z8>
- Twenge, J. M., & Campbell, W. K. (2018). Associations between screen time and lower psychological well-being among children and adolescents: Evidence from a population-based study. *Preventive Medicine Reports*, 12, 271-283. <https://doi.org/10.1016/j.pmedr.2018.10.003>
- UNESCO. (2024). *Informe de seguimiento de la educación en el mundo, 2023: Tecnología en la educación: ¿una herramienta en los términos de quién?* GEM Report UNESCO. <https://doi.org/10.54676/NEDS2300>
- Vaid, S. S., Kroencke, L., Roshanaei, M., Talaifar, S., Hancock, J. T., Back, M. D., Gosling, S. D., Ram, N., & Harari, G. M. (2024). Variation in social media sensitivity across people and contexts. *Scientific Reports*, 14(1), 6571. <https://doi.org/10.1038/s41598-024-55064-y>
- Valkenburg, P. M. (2022). Social media use and well-being: What we know and what we need to know. *Current Opinion in Psychology*, 45, 101294. <https://doi.org/10.1016/j.copsy.2021.12.006>
- Valkenburg, P. M., Beyens, I., & Keijsers, L. (2022). *Investigating Heterogeneity in (Social) Media Effects: Experience-Based Recommendations*. <https://doi.org/10.31234/osf.io/rwc73>
- Vogel, E. A., Rose, J. P., Roberts, L. R., & Eckles, K. (2014). Social comparison, social media, and self-esteem. *Psychology of Popular Media Culture*, 3(4), 206-222. <https://doi.org/10.1037/ppm0000047>



- Wu, T. (2016). *The Attention Merchants: The Epic Scramble to Get Inside Our Heads*. *Faculty Books*. <https://scholarship.law.columbia.edu/books/64>
- Zhang, S., Zhao, X., Zhou, T., & Kim, J. H. (2024). Do you have AI dependency? The roles of academic self-efficacy, academic stress, and performance expectations on problematic AI usage behavior. *International Journal of Educational Technology in Higher Education*, 21(1), 34. <https://doi.org/10.1186/s41239-024-00467-0>
- Zhang, W., Pu, J., He, R., Yu, M., Xu, L., He, X., Chen, Z., Gan, Z., Liu, K., Tan, Y., & Xiang, B. (2022). Demographic characteristics, family environment and psychosocial factors affecting internet addiction in Chinese adolescents. *Journal of Affective Disorders*, 315, 130-138. <https://doi.org/10.1016/j.jad.2022.07.053>
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Profile books.

