

UNIVERSIDAD COMPLUTENSE DE MADRID  
FACULTAD DE FILOSOFÍA

---

MÁSTER UNIVERSITARIO EN  
ÉTICAS APLICADAS

Trabajo de Investigación

# PARASITISMO LIBERTARIO EN ENTORNOS DIGITALES

Análisis del problema y propuestas de actuación

Luis Diego SARABIA TORRES

TUTOR: Prof. Dr. Juan Antonio Fernández Manzano

Madrid, septiembre de 2021

# Índice de contenido

Resumen.....	3
Abstract.....	3
Palabras clave / Keywords.....	3
1. INTRODUCCIÓN.....	4
2. EL PATERNALISMO LIBERTARIO.....	8
2.1 ¿Cabe pensar un parasitismo libertario?.....	9
3. ECONOMÍA DE LA ATENCIÓN.....	11
4. PROBLEMAS.....	16
4.1 Irrealidad, posverdad y polarización.....	19
4.1.1 Filtro burbuja.....	24
4.1.2 El medio es el mensaje.....	26
4.2 Irrealidad, adicción y trastornos.....	28
5. ¿POR QUÉ FUNCIONAN?.....	32
5.1 Racionalidad limitada (bounded rationality).....	33
5.2 La teoría argumentativa o interaccionista de la razón.....	34
5.3 A/B Testing.....	36
5.4 La banalidad del mal digital: La convergencia instrumental y extremismo algorítmico.....	37
6. LA CUESTIÓN DE LA RESPONSABILIDAD INDIVIDUAL Y EL CASO DE LA ADICCIÓN .....	38
6.1 El individuo frente a la estructura.....	40
7. LA VÍA DE LA RESPONSABILIDAD SOCIAL DE LAS EMPRESAS.....	42
8. PROPUESTAS.....	44
8.1 Propuestas de aumento de la información.....	44
8.2 Propuestas desde el paternalismo libertario.....	45
8.3 Cambios al incentivo de los algoritmos.....	46
8.4 Cambios legales significativos.....	47
9. CONCLUSIONES.....	49
Referencias bibliográficas.....	51

## **Resumen**

El presente trabajo trata de mostrar el hilo conductor que une algunos importantes problemas de la actualidad a través de su relación con la arquitectura de las redes sociales. Estos problemas se presentan en dos categorías principales: sociales (posverdad, polarización) e individuales (adicciones, ansiedad, depresión). Las redes sociales generarían o, como poco, potenciarían estos problemas de forma coherente con lo que se denominará parasitismo libertario, es decir, dirigiendo a los individuos hacia comportamientos que van en su propio perjuicio, por medio de un sistema de incentivos propiciado por el modelo económico de la economía de la atención, pero siempre bajo la premisa de que el individuo es formalmente libre para dejar de usarlas. En este modelo, las empresas obtienen más beneficios cuanto más tiempo consigan mantener a más usuarios dentro de sus redes y produzcan mejores modelos predictivos con los que vender espacios a anunciantes. Se apuntarán algunos de los mecanismos mediante los que el diseño de estas redes sociales y sus algoritmos se convierten en algo potencialmente perjudicial y cómo su defensa proviene de una antropología deficiente. También se discutirá la posibilidad de una solución desde dentro del propio entorno empresarial a través del concepto de responsabilidad social de las empresas y, por último, se bosquejarán algunas propuestas de solución.

## **Abstract**

The present work intends to show the common thread that links some important current problems through their relationship with social network's architecture. These problems fall into two main categories: social (post-truth, polarization) and individual (addictions, anxiety, depression). Social networks would generate or, at the very least, augment these problems in a way consistent with what will be called libertarian parasitism, that is, directing individuals towards behaviors that are to their own detriment, due to a system of incentives set by the economic model of the attention economy, but always under the premise that the individual is formally free to stop using them. In this model, companies profit more the longer they manage to keep more users within their networks and produce more accurate predictions of their behaviour in order to get ad revenue. It will also outline some of the mechanisms by which the design of these social networks and their algorithms become potentially harmful and how its advocates uphold an inaccurate anthropology. It will also explore the possibility of a solution from within the business environment itself through the concept of corporate social responsibility and, finally, it will sketch some proposals to fix the problems generated.

## **Palabras clave / Keywords**

Redes sociales, economía de la atención, paternalismo libertario, adicción, posverdad

Social networks, attention economy, libertarian paternalism, addiction, post-truth

## 1. INTRODUCCIÓN

*El verdadero problema de la humanidad es el siguiente: tenemos emociones paleolíticas, instituciones medievales y tecnología de dioses.*

Edward O. Wilson

El 6 de enero de 2021 tenía lugar en el Congreso de EE.UU. la sesión en la que se debían certificar los resultados de las elecciones presidenciales del 3 de noviembre, en la que Biden se impuso sobre Trump. Esta formalidad procedimental venía precedida de polémica, ya que muchos de los seguidores del aún presidente exigían a los republicanos que suspendieran la cuenta de los votos de los colegios electorales, en medio de acusaciones de fraude masivo en los comicios. Hasta seis senadores republicanos presentaron objeciones al recuento debido a esas alegaciones, aunque los restantes 93 miembros de la Cámara Alta votaron en contra de la objeción (Steven et al., 2021) y el vicepresidente Pence publicó una carta desmarcándose de la posibilidad de impedir el recuento. Hasta el momento, todas las demandas judiciales interpuestas en los diferentes estados en disputa, además del Tribunal Supremo, han negado la existencia de tal fraude.

Mientras esto ocurría, en una comparecencia pública, Trump y Giuliani arengaban a sus seguidores a bajar por la avenida Pennsylvania para parar lo que consideraban un robo, una ilegalidad, unas elecciones menos honestas que las de países del tercer mundo, etc. (Kambhampati et al., 2021). Muchos de los asistentes se pusieron en marcha y asaltaron el Capitolio, comenzando disturbios, obligando a suspender la sesión y protagonizando hechos e imágenes históricamente anómalos que desembocaron en la muerte de varias personas y peleas en las calles de Washington. Mientras ocurrían los hechos, Trump pidió por Twitter a sus seguidores que se fueran a casa, a la vez que justificaba el asalto; decía que los asaltantes eran queridos y especiales, insistía en el robo de unas elecciones en las que él supuestamente habría arrasado y acusaba de cobardía a su propio vicepresidente. Horas más tarde, Twitter reaccionaba y eliminaba tres de los tweets del presidente<sup>1</sup>, que antes había marcado como «en disputa», como resultado de la situación violenta y la violación de las políticas de Integridad Cívica de la red social<sup>2</sup> y suspendía su cuenta. Más tarde lo harían Facebook e Instagram<sup>3</sup>, seguidos de multitud de otras redes sociales (Denham, 2021)<sup>4</sup>. El conflicto corporativo siguió escalando de nivel cuando grandes cantidades de partidarios de Trump y personas afines en otros países se pasaron a la aplicación Parler en nombre de la libertad de expresión, pero esta fue retirada de los servidores de Amazon, Apple y Google por incumplir normas del servicio (Fung, 2021).

1 <https://media-cdn.factba.se/realdonaldtrump-twitter/1346900434540240897.jpg>  
<https://media-cdn.factba.se/realdonaldtrump-twitter/1346928882595885058.jpg>  
<https://media-cdn.factba.se/realdonaldtrump-twitter/1346954970910707712.jpg>

2 <https://twitter.com/TwitterSafety/status/1346970430062485505>

3 <https://twitter.com/fbnewsroom/status/1346994037786619905?s=20>

4 Entre las importantes: Youtube, Snapchat, Reddit y Twitch.

Durante la polémica en las calles y en las diferentes redes sociales, muchos seguidores de Trump acusaban a los políticos republicanos que acataron las elecciones y a las grandes corporaciones de fascismo, traición a la patria, comisión de todo tipo de atropellos ilegales y, simultáneamente, otros declaraban que los asaltantes eran secretamente miembros de Black Lives Matter o Antifa, rumores alentados en algunos casos por representantes republicanos electos<sup>5</sup>. En medio de esta guerra de desinformación, algunos medios de verificación intentaron establecer qué estaba pasando (Jacobson et al., 2021), pero se veían desbordados. El fenómeno es llamativo en la medida en la que parece apuntar a que algunas personas viven una realidad completamente ficticia, sean los seguidores de Trump o quizá el resto del mundo. Mitch McConnell, líder de la mayoría republicana en el Senado, en su declaración ante el resto de senadores poco antes del asalto, había rechazado las teorías de la conspiración, las alegaciones de fraude y declarado que si las elecciones se anulaban en base a meras alegaciones del bando perdedor, la democracia entraría en una espiral de muerte. Llamativamente, añadió<sup>6</sup>:

El autogobierno requiere un compromiso compartido con la verdad. Y un respeto compartido por las reglas básicas de nuestro sistema. No podemos seguir dividiéndonos en dos tribus separadas con un conjunto separado de hechos y realidades separadas sin nada en común excepto nuestra hostilidad hacia los demás y la desconfianza hacia las pocas instituciones nacionales que todos compartimos todavía.

Meses antes, en Naperville, una pequeña ciudad de Illinois, Alexander E. Kearns se quitaba la vida dejando una breve nota que decía «¿Cómo pudo un joven de 20 años sin ingresos tener asignado casi un millón de dólares de apalancamiento?» (Klebnikov, 2020). Kearns estaba haciendo referencia al margen del que había dispuesto en una aplicación de inversión en bolsa llamada Robinhood, que ofrece la posibilidad de comerciar con acciones de una forma divertida, sencilla y sin comisiones, obsequia a sus nuevos usuarios con acciones gratuitas al comienzo y dispone de un sistema de recompensas, niveles y objetivos que fomentan su uso. Kearns, sin saber apenas lo que estaba haciendo, había llegado a contraer una deuda de 730,165 dólares<sup>7</sup>, aunque dejó escrito que jamás tuvo la intención de arriesgarse tanto. Mediante la gamificación, Robinhood atrae y mantiene a inversores jóvenes que terminan por perderlo todo, aunque estos anteriormente hubieran sido inversores ocasionales y prudentes con las formas tradicionales de inversión. Mediante «empujoncitos» (*behavioural nudges*) como el interfaz divertido, las transiciones con emojis, el confetti virtual cuando se realizan inversiones, las notificaciones constantes que recuerdan las subidas, bajadas y otros productos, así como los puntos y recompensas que se obtienen según se alcanzan ciertos objetivos dentro de la aplicación, atrapan a muchos usuarios lo suficiente como para que arriesguen su dinero más de lo que lo harían en otras

---

5 <https://twitter.com/RepMoBrooks/status/1346899440989310980>

6 Traducción propia. Transcripción del discurso en Phillips (2021).

7 Si bien, parece ser que no comprendió algunas de las complejidades de mercadear con opciones y que su deuda se hubiera corregido en pocos días

condiciones, perdiendo el apetito y el sueño además de su dinero a medida que pasan más y más tiempo en la aplicación (Popper, 2020). De hecho, cuanto más tiempo pasan en ella, más probable es que se arruinen (Choy, 2015 y Barber et al., 2014). Sin embargo, esto es esencial para el rendimiento de la aplicación, pues el portal obtiene más dinero si los clientes compran y venden más, de modo que la aplicación tiene un incentivo inapelable para fomentar este tipo de uso abusivo. Los clientes de Robinhood son, estadísticamente y por un amplio margen, los que negocian en productos de más riesgo y a mayor velocidad.

También han aumentado espectacularmente los suicidios, especialmente entre las jóvenes, al hilo de la escalada de trastornos de la ansiedad y de depresión en adolescentes desde 2003 a la actualidad (de 5,4% a 20%, datos de EE.UU, Muiruri, 2021). Este crecimiento coincide con el aumento de tiempo de pantalla desde la llegada de los smartphones y las redes sociales. Estudios estadísticos han analizado la correlación entre ese tiempo de pantalla y las tasas de ansiedad y depresión (así como con el uso intensivo de videojuegos o recibir cyberbullying, factores que ya se habían estudiado) mostrando una correlación clara (Muiruri, 2021 y Heid, 2019) que, como veremos, Twenge (2017) es partidaria de circunscribirla más bien a las redes sociales que a otros usos del tiempo de pantallas. Las nuevas generaciones están más deprimidas y ansiosas que las anteriores. Su vida transcurre fluidamente entre el mundo físico y el digital de forma cotidiana y no es extraño que incluso antes de la adolescencia pasen decenas de horas a la semana con sus dispositivos, especialmente en TikTok, Instagram y Youtube<sup>8</sup>. La relación entre el tiempo en pantalla y los desórdenes mentales es sutil y compleja de explicar o probar en general, pero en otros casos resulta más evidente, como por ejemplo el hecho de que durante años el algoritmo de recomendación de Youtube sugirió cientos de millones de veces, de forma sistemática, canales como «thinspiration», relacionados con la exaltación de la anorexia, a adolescentes que veían vídeos sobre dietas (Harris, 2021).

Las redes sociales también afectan a la salud de las personas mediante la desinformación, por ejemplo mediante la propagación de información contra las vacunas, que está multiplicando los casos de sarampión y otras enfermedades en muchas partes del mundo, especialmente las más desfavorecidas (UN News / Health, 2019) en las que no existen estructuras educativas o periodísticas estables y respetadas, por lo que la gente no tiene suficientes elementos externos con los que contrastar la desinformación que reciben de las redes sociales y puede llegar a creer sinceramente que las medicinas o las vacunas son métodos de control o exterminio de la población. Quienes generan estos contenidos consiguen viralidad aprovechando el miedo legítimo que tienen los padres y madres a inyectar a sus

---

8 En una investigación propia llevada a cabo con dos grupos de 25 estudiantes madrileños de 1º de la ESO (edades entre 11 y 13 años) en el curso 20/21, en al menos un 10% de los casos el uso semanal superaba las 40 horas. Estos datos son compatibles con la información divulgada por Pantallas Amigas de su estudio SmartHuman, que indica que el 40% de los alumnos de 1º de la ESO lo utilizan entre 1 y 3 horas, el 18% entre 3 y 6 horas. En las franjas de uso de 6 a 9 horas se encuentra el 6% del alumnado, mientras que en la de 9 a 12 y más de 12 están el 3% y el 4% respectivamente. Divulgado en: <https://www.pantallasamigas.net/estudio-smarthuman-concienciacion-uso-movil-adolescentes-adicciones-digitales/>. En mi investigación no encontré a ningún caso en esas franjas de uso.

hijos algo que podría matarlos.

El aumento de la polarización, el fomento de las adicciones o el riesgo financiero, el aumento de los suicidios y los trastornos mentales, etc., nos obligan a plantear cuestiones filosófica y políticamente fundamentales con respecto a la comunicación, el ser humano, las nuevas tecnologías y la democracia.

- ¿Existe algún hilo común que conecte estos problemas (políticos, sociales, individuales, médicos, económicos, etc.) pese a su aparente disparidad?
- ¿En qué medida los algoritmos que rigen los flujos de información tienen sesgos o mecanismos peligrosos para la sociedad en general o los individuos en particular? ¿Qué tipo de relación hay entre las plataformas de los gigantes tecnológicos y problemas como la adicción, la depresión, la ansiedad, la posverdad, la polarización, etc.?
- ¿Por qué caemos en estos problemas? ¿Qué tipo de racionalidad posee el ser humano? ¿Existen sesgos cognitivos y emocionales que puedan ser explotados para fines interesados? ¿De quién es la responsabilidad sobre estos problemas?
- ¿Este problema debe dejarse a la autorregulación de las empresas? ¿Pertenece al ámbito responsabilidad social de las empresas? ¿De quién es la responsabilidad cuando ocurren acontecimientos en el mundo real que fueron provocados o potenciados por eventos que ocurrieron en el mundo virtual? ¿Existen medidas que quepa exigir ética o legalmente en el diseño de las Plataformas *Online* Basadas en la Economía de la Atención<sup>9</sup> (POBEA en adelante)?

El abordaje de estas cuestiones se ha realizado a través de la revisión de literatura, tanto específicamente filosófica como académica en general. Dado que se trata de una temática de actualidad, la mayor parte de la bibliografía utilizada es muy reciente, por lo que existen amplios debates sobre muchos de los problemas tratados. Además, debido a la relación con la realidad política y social presente que tienen estos problemas, a menudo se ha recurrido a fuentes periodísticas, revistas especializadas y material más divulgativo para documentar los hechos, problemas o reacciones de los diferentes agentes implicados.

La hipótesis de partida de este trabajo es que estos problemas, con todas sus diferencias, comparten algunas raíces comunes al uso de las POBEA producto de un sistema de incentivos no centrado en el

---

<sup>9</sup> Mediante el término plataformas *online* basadas en la economía de la atención se quiere englobar a todas aquellas plataformas que funcionan según lo explicado en el apartado 3 y generan los problemas del apartado 4. Estas plataformas son principalmente redes sociales. Sin embargo, limitar el análisis a las redes sociales sería insuficiente, ya que muchas de las principales prácticas y consecuencias que se mostrarán también se dan a través de otro tipo de portales que no caen fácilmente en esa categoría. Youtube, Netflix o Google, son importantes en la recopilación masiva de datos, la creación de filtros burbuja y modelos predictivos individualizados o la segmentación de los contenidos. La ya mencionada app Robinhood tampoco es una red social, pero comparte con ellas características de explotación de la economía de la atención, como la ludificación, las notificaciones, etc. La enorme variabilidad de las formas en las que se presentan estos fenómenos y la ausencia de un término común específico no implica la ausencia de líneas comunes (y, con ello, posibles soluciones comunes), por lo que se utilizará el acrónimo POBEA de manera general.

interés de los usuarios y, a través de sus mecanismos, son responsables de empujar a sus usuarios hacia comportamientos que van en su propio perjuicio, si bien siempre bajo la premisa de que el individuo las utiliza por propia elección, en lo que se denominará parasitismo libertario. Sus efectos no se limitan a la población que habita los mundos digitales, ya que, aunque no se usen redes sociales, todas las personas viven rodeadas de otras que lo hacen, y estas modelan su forma de pensar, comportarse, desear, etc. Estar al margen de las redes sociales no evita estos problemas, aunque quizá sí mitigue los más específicamente individuales, porque convivimos en entornos en los que nuestros vecinos pueden estar radicalizándose en redes y votando o actuando en consecuencia, enviando a sus hijos sin vacunar al colegio, siendo objeto de propaganda que conduce a desórdenes de la alimentación o depresiones, etc.

## **2. EL PATERNALISMO LIBERTARIO**

Para comprender qué es el parasitismo libertario, conviene primero explicitar de qué idea es su opuesto. El paternalismo libertario (Thaler y Sunstein, 2003) es una propuesta de intervención social, especialmente en el entorno de las instituciones públicas, pero no necesariamente limitado a ellas, que sostiene que es posible y deseable ejercer una influencia en el comportamiento de los ciudadanos en su propio beneficio respetando, simultáneamente, la libertad de decisión de los individuos. En muchas de estas situaciones, ni siquiera hay una posición neutra que pudieran adoptar las instituciones, sino que su decisión, en cualquier sentido, ya será una toma de partido. Un ejemplo típico sería la cuestión de la donación de órganos. En algunos países, la donación de órganos requiere de un acto voluntario del individuo para apuntarse a una lista, mientras que en otros, los individuos están apuntados por defecto y tienen la opción de borrarlos si así lo desean. Ninguna de las dos opciones es más neutral que la otra, pero la segunda obtiene unas tasas de donación de órganos espectacularmente superiores, con la consiguiente mejora de la salud pública. El paternalismo libertario apostaría por medidas de este tipo que guíen las elecciones de los individuos hacia su propio interés (o, como en el caso de los trasplantes, el interés social), sin menoscabar su libertad de elección, ya que en todo momento pueden optar por alternativas. El diseño de las instituciones, los espacios comerciales, los mecanismos de acceso, las normas, etc., tiene una influencia enorme en las elecciones que hacen los individuos y, por lo tanto, es una buena idea que estén orientadas hacia el bienestar individual y social. En muchas decisiones, además, los individuos carecen de preferencias claras, estables o bien orientadas y sus elecciones se encuentran determinadas por motivos no racionales, como seguir la regla tradicional, los efectos marco o los puntos de partida, etc. (Sunstein y Thaler, 2003, pp. 9-12). Así, se deberían diseñar las opciones de tal manera que, por defecto, se tome la opción preferida por el planificador, pero asegurando que los costos de salirse de ella sean triviales. Otros ejemplos serían la participación por defecto en planes de pensiones con las nóminas de los trabajadores o la colocación de la comida más saludable en la zona más visible de una cantina. Los individuos siempre podrían salirse del plan de pensiones con una rápida gestión o caminar unos pasos para optar por comida menos saludable, pero quienes tomen decisiones poco motivadas, tenderán a hacer lo que es mejor para ellos. Este paternalismo dejaría de ser libertario



si se obliga a participar en el plan de pensiones, se prohíbe completamente la comida menos saludable o se la sitúa a 300 metros de distancia.

El paternalismo libertario es deseable en gran cantidad de situaciones, ya que los planificadores tienen que establecer puntos de partida o reglas en las instituciones o situaciones que están diseñando necesariamente, y, dado que tendrán que tomar esas decisiones, deberían estar orientadas al bienestar de los individuos en condiciones de garantía de la máxima autonomía posible para aquellos que quieran optar por alternativas<sup>10</sup>. Thaler y Sunstein proponen diferentes escenarios, cada vez menos libertarios y más paternalistas, en función del contexto y las restricciones que imponen la racionalidad y el autocontrol limitados de los individuos: un paternalismo mínimo (el que ocurre cuando hay que diseñar necesariamente un punto de partida: una política, institución, etc.), las elecciones activas requeridas (forzar a la gente a elegir, sin predeterminedar ninguna opción), las restricciones de procedimiento (situaciones en las que salirse de la opción planificada viene acompañado de ciertas restricciones, para asegurarse que el que opta por otra cosa lo hace por voluntad propia, por ejemplo, forzar a la gente a llevar a cabo un proceso administrativo para renunciar a algo), restricciones sustantivas (en la que la salida del modelo planificado solo puede darse en condiciones decididas por los planificadores) y el rechazo de la libertad de elección (cuando los que rechacen el plan tradicional se equivocarán siempre o la mayor parte de las veces, por ejemplo, cerrar un puerto de montaña durante una nevada). Dadas estas categorías, un paternalista libertario buscaría maximizar las situaciones en las que se dan las dos primeras opciones o, excepcionalmente la tercera y la cuarta, siempre supeditadas a estudios de costos y beneficios o siguiendo algún tipo de razonamiento que lo justifique.

Esta teoría parte de una crítica al modelo de racionalidad de la escuela neoclásica de economía, según la cual el ser humano se comportaría de forma racional ante la información que recibe y la procesaría adecuadamente, buscando siempre el nivel más alto de bienestar y satisfacción de sus propios intereses dadas las oportunidades y restricciones de su entorno. Frente a este modelo, el paternalismo libertario parte de una teoría de la racionalidad limitada (*bounded rationality*), que trataremos más adelante y que defiende que nunca contamos con la información suficiente, que nuestras habilidades cognitivas son limitadas y que poseemos un autocontrol también limitado. Dado que en muchos contextos los individuos carecen de la inteligencia, experiencia o información suficiente para tomar buenas decisiones, es adecuado que existan políticas que orienten sus decisiones. Las elecciones de los individuos, a menudo, son cuestionables, incluso para sí mismos. Un ejemplo sería la obesidad, causante de miles de muertos anualmente y provocada por malas decisiones con respecto a la dieta. En estos contextos, las decisiones de los individuos no son la mejor manera de promover su propio bienestar.

## **2.1 ¿Cabe pensar un parasitismo libertario?**

---

10 De ahí el término paternalismo (actuar por el bien/protección de otros) libertario (ofreciendo máxima autonomía individual)

Si existe el paternalismo libertario como propuesta de guiar a los individuos para la consecución de sus intereses permitiendo la libertad de tomar decisiones al margen del planificador, ¿podría existir su inverso? Es decir, ¿existen situaciones en las que los individuos, aunque formalmente libres, estén planificadamente guiados hacia situaciones que perjudican su bienestar pero que benefician al arquitecto de la decisión? Y, en caso de existir y suponiendo que no exista una motivación de maldad moral, ¿cuáles serían los incentivos que generan este tipo de situaciones?

En este trabajo se defenderá que existen tales situaciones en diferentes ámbitos y que dichas situaciones son las soluciones más racionales posibles dado el tipo de entorno en el que existen, pese a que tengan efectos muy perjudiciales. El ejemplo más ilustrativo del parasitismo libertario, lo tenemos en los casinos. Los casinos desarrollan su actividad en un modelo de negocio que sitúa sus intereses en contra de los de sus clientes. Por supuesto, un casino puede aportar muchos servicios, como comida y bebida, espectáculos, etc., pero su principal fuente de ingresos reside en los juegos de azar donde los clientes tienen una ligera desventaja estadística frente a la casa. En principio, cuanto más tiempo pasen apostando los clientes, más probable es que gane la banca y, aunque existen jugadores profesionales que hacen cierto dinero, existen todavía más perdedores (más o menos habituales) que lo compensan. Los casinos no obligan a nadie a jugar, ni impiden la salida de sus instalaciones a aquellos que han ganado dinero y solo muy ocasionalmente vetan la entrada a los que ganan (García-Pelayo, 2004), algo ilegal, por otra parte (Hernández, 2000). Sin embargo, los casinos despliegan toda una batería de estrategias para forzar a los clientes contra sus propios intereses. Las más evidentes serían las bebidas alcohólicas gratuitas a los jugadores con más dinero o la recompensa con noches gratis en buenas habitaciones a aquellos que se han dejado mucho dinero en anteriores ocasiones. Sin embargo, hay muchas otras estrategias más sutiles, relacionadas con la arquitectura, el diseño de los espacios, luces, sonidos y máquinas, ideadas para reducir la agencia de los clientes y maximizar el tiempo que pasan en el casino.

Como explica Schüll, (2012, p. 5), las grandes protagonistas de los casinos no son ya las mesas de blackjack o la ruleta, al menos no en cuanto a los beneficios se refiere, sino las máquinas tragaperras. En 2003, más del 85% de los beneficios de los casinos de Nevada provenían de ellas. Una máquina estándar es un complejo dispositivo, generalmente digital (aunque aún existen las que se componen de partes mecánicas), hecho de miles de piezas individuales mediante un proceso que involucra a aproximadamente trescientas personas, entre guionistas, diseñadores gráficos, matemáticos, agentes de marketing, ingenieros electrónicos y de software y diseñadores de sistemas subsidiarios como pantallas táctiles, validadoras de billetes o fichas NFC, etc. Los clientes se enfrentan a un dispositivo calculado, diseñado y perfeccionado a lo largo de décadas e innumerables estudios de mercado, colocado ya en un espacio planificado cuidadosamente para evitar sobresaltos y eliminar cualquier tipo de señales de parada (*stopping cues*)<sup>11</sup>. El diseño de las máquinas y los casinos en los que se alojan ha llegado a ser tan

11 Una señal de parada o *stopping cue*, es toda señal que obliga al individuo a pararse y tomar una decisión que puede ser, potencialmente, hacer algo nuevo. En los casinos, la supresión más clara es la eliminación de las ventanas, que hace que los individuos no perciban los ciclos horarios (la caída de la luz o el amanecer puede ser una señal de

absorbente que los usuarios llegan a no darse cuenta de que un escape de agua les ha sumergido los pies completamente, que han saltado las alarmas de incendio o que la persona de la máquina de su lado está siendo reanimada con un desfibrilador.

Pese a la enorme asimetría de poder que ejercen los casinos sobre sus clientes, estos son libres, en última instancia, de levantarse y salir. ¿O no? ¿Es posible que el conocimiento de la naturaleza humana acumulado durante décadas de aprendizaje acerca de los clientes haya sobrepasado el umbral de resistencia de quienes entran en un casino? ¿Los mecanismos que despliegan pueden llegar a hacer perder el control de sus acciones a los clientes? Incluso si no ha llegado aún ese punto más que para las personas de bajo autocontrol y con tendencia a la ludopatía<sup>12</sup>, ¿no existe un incentivo claro en su modelo de negocio para que los casinos profundicen en las técnicas que despliegan actualmente, hasta hacerlas todo lo seductoras y adictivas que sea posible? ¿Qué hacer con estas situaciones de parasitismo libertario<sup>13</sup>? ¿Qué elementos de diseño las fomentan? ¿Qué propuestas podrían hacerse para limitarlas?

En el presente trabajo, se estudiará principalmente el problema desde la perspectiva de las POBEA, pero la discusión teórica y algunas de las soluciones prácticas se podrían aplicar a otros contextos como el de los casinos. Ya que la esencia del modelo de negocio es esencialmente la misma: ni los casinos, ni las redes sociales en particular ni las POBEA en general ganan dinero vendiendo directamente un producto, sino manteniendo al usuario en ellas. A su modelo se le ha llamado economía de la atención.

### 3. ECONOMÍA DE LA ATENCIÓN

Las POBEA no son un producto que compren sus usuarios, ya que el acceso a ellas es gratuito. Su complejidad de diseño y el mantenimiento que requieren se ofrecen a coste cero a los miles de millones de personas que las utilizan de forma regular. Sin embargo, son negocios tan gigantescos que la capitalización de mercado de las empresas que las gestionan supera al PIB de países enteros. Hasta 5 de ellas se encuentran entre las 25 economías más grandes del mundo: Apple (8ª), Microsoft (10ª), Amazon (15ª), Alphabet (16ª) y Facebook (22ª)<sup>14</sup>. Por comparar, la única compañía de otro sector que se

---

parada). Otra más sutil es la reducción de los ángulos en el diseño de los pasillos. Un ángulo supone una señal de parada, mientras que una curva suave aumenta la probabilidad de que el cliente vaya a donde se le dirige. En una revista, el final del artículo es una señal de parada, en las redes sociales, las pantallas de despliegue infinito (se usará el término técnico en inglés - *infinite scrolling feeds* - en adelante) las eliminan. Lo explica Adam Alter en el minuto 5 de [https://www.ted.com/talks/adam\\_alter\\_why\\_our\\_screens\\_make\\_us\\_less\\_happy/](https://www.ted.com/talks/adam_alter_why_our_screens_make_us_less_happy/)

12 Cabe destacar que Schüll defiende taxativamente lo contrario, el diseño de los casinos y las máquinas tragaperras no atrae a un modelo de ludópata adicto y disfuncional que posee algún tipo de problema de carácter, sino que los mecanismos que despliegan son análogos a lo que hacen ciertas drogas con el cerebro humano y sobrepasan la capacidad de resistencia del individuo medio. En la misma línea se sitúa Alter (2018), quien también considera errónea la tendencia a pensar en la adicción como algo inherente a algunas personas, como si hubiera un perfil psicológico del «adicto». Lo que se sabe de la naturaleza general de la adicción a día de hoy apunta a que las causas más importantes para propiciar las adicciones son las circunstancias y el entorno de la persona.

13 Es de esperar que todos aquellos preocupados por la autonomía de los sujetos, incluyendo a los que juzgan con severidad el paternalismo libertario, consideren que esta es una cuestión relevante y que si aterra pensar en las pendientes resbaladizas y los errores de planificación del Estado cuando vela por nuestros intereses, también habrá que preocuparse por las pendientes resbaladizas y, en este caso, los aciertos de planificación cuando los intereses privados juegan en contra de los de los individuos.

14 Comparando capitalización de mercado vs PIB. <https://www.visualcapitalist.com/the-tech-giants-worth-compared->

encuentra en esa lista es Saudi Aramco, dedicada al comercio de petróleo, gas y productos derivados, y que gestiona el 10% del petróleo mundial.

Para llegar a tal tamaño económico, estas compañías necesitan escala (cantidad de usuarios) y compromiso (*engagement*) de sus usuarios, que son valiosos para las POBEA porque son valiosos para los anunciantes. Así, lo que hacen estas compañías es vender la atención de los usuarios. Los verdaderos clientes, los que pagan por los servicios de estas empresas, no son los usuarios que pasan tiempo en ellas, sino los anunciantes que utilizan esos espacios para atraer a esos usuarios hacia sus productos. Pero si su funcionamiento es similar al de televisión o radio, ¿por qué se han alcanzado tales cotas de capitalización de mercado y no las productoras televisivas, que ya llevan décadas en el mercado? Porque estos gigantes tecnológicos han encontrado la forma de vender algo más: certeza. En la televisión, la radio o los periódicos, los anuncios y los mensajes se segmentan en función de características genéricas que se conocen sobre sus usuarios a través de estudios de mercado. Sin embargo, las redes sociales y demás aplicaciones pueden recolectar datos personalizados de cada cliente, crear perfiles predictivos y segmentar la publicidad de forma personalizada. Se acumulan montañas de datos extraídos de cada uno de los usuarios, pero no para los usuarios, sino para generar modelos de esos usuarios con los que poder ofrecer un mayor grado de certeza predictiva a los anunciantes. Conociendo a los clientes, se puede segmentar la publicidad de tal manera que se maximice la probabilidad de que el usuario realice la conducta que se desea de este, sea comprar un producto, votar a un partido político, apuntarse a un grupo de reciclaje o unirse a una acción política. El valor de las compañías tecnológicas está en los datos que puedan recoger y los modelos algorítmicos con los que predecir la conducta de los usuarios. Así, aportan un mayor grado de certeza a los anunciantes (empresas, partidos políticos, ONGs, agentes de desinformación, quien pague) de que, bien segmentado, su anuncio será capaz de modificar el comportamiento de los usuarios con un grado de éxito más elevado.

La predicción de la conducta, esencial para vender certeza, se da de dos modos fundamentales:

1. Construyendo mejores modelos predictivos del ser humano, más precisos, más capaces, a través de la acumulación de datos y la experimentación.
2. Modificando la conducta<sup>15</sup>. Fomentando progresivamente ciertos modelos de comportamiento humano más simples, más predecibles. Es decir, si se educa sistemáticamente al público a realizar ciertos tipos de comportamiento, a la búsqueda de dopamina, a la reacción basada en

---

[economies-countries/](#) Datos a 13 de junio de 2021. Se puede cuestionar la presencia de Apple en esta lista, ya que ha tendido a evitar la explotación de la economía de la atención (Zuboff, pp. 38-39).

15 Zuboff (2020, 18): «Con el tiempo, los capitalistas de la vigilancia descubrieron que los datos conductuales más predictivos se obtienen interviniendo en la marcha misma de las cosas para empujar a, persuadir de, afinar y estimular ciertos comportamientos a fin de dirigirlos hacia unos resultados rentables.» El desarrollo de los mecanismos mediante los que se hace esto los trata Zuboff en el capítulo 10 y describe principalmente tres: afinación, arreo y condicionamiento. El escritor e informático Jaron Lanier, de hecho, propone llamar a Google y Facebook imperios de modificación de conducta (*behavior-modification empires*): [https://www.ted.com/talks/jaron\\_lanier\\_how\\_we\\_need\\_to\\_remake\\_the\\_internet](https://www.ted.com/talks/jaron_lanier_how_we_need_to_remake_the_internet)

respuestas de la amígdala, etc., se convierte a los usuarios en personas más reactivas, más indignables, más preocupadas por su estatus personal y la percepción social que hay de ellos, más narcisistas, más adictas a la atención<sup>16</sup>.

De este modo, de lo que se trata es de expandirse, aumentar el compromiso y conocer el comportamiento humano, predecirlo, darle forma y orientarlo hacia los fines de otros. El nivel de predicción es tal que Facebook incluso vende a los anunciantes cuándo un usuario va a dejar de ser leal a una marca (*Facebook loyalty prediction*) (Hern, 2020)<sup>17</sup>, y el objetivo final ideal sería el logro de modelos predictivos que se aproximen a la certeza absoluta.

También es fundamental, como ya se mencionó, la escala (cantidad de usuarios). Aquí, sin embargo, parecería que nos topamos con algún límite antropológico-cognitivo que impide el desarrollo de los objetivos de las POBEA. Por ejemplo, el antropólogo Robin Dunbar (1998) sostiene que existe una relación entre el tamaño del cerebro de los primates y la media del tamaño de los grupos sociales que forman. Un individuo determinado puede mantener de manera cómoda en torno a 150 relaciones estables de amistad, conocido como número de Dunbar y definido informalmente por este como «la cantidad de personas a las que uno no se sentiría avergonzado de unirse sin invitación si se los encontrase en un bar» (1998, 77)<sup>18</sup>. Esta frontera está limitada por nuestras capacidades cognitivas y por el tiempo que requiere atender estas relaciones. Por otra parte, el éxito de las redes sociales está determinado por el número de usuarios y la cantidad de conexiones entre ellos, ya que, como se ha mostrado en Shriver, Nair y Hofstetter, (2013), el aumento de conexiones e interacciones entre usuarios genera un incentivo psicológico para que estos aumenten la cantidad de contenido que producen, aumentando así el tiempo que otros usuarios pasan en la plataforma y creando contenido, lo que produce un efecto multiplicador. De este modo, hay una relación entre la creación y refuerzo de lazos sociales y la monetización de la red social<sup>19</sup>. Pero entonces, esto significaría que hay un tope impuesto por los límites propios de nuestra especie al tamaño del negocio que una red social puede esperar conseguir. Sin embargo, lo que hacen estas es utilizar mecanismos que suponen una suerte de «dopaje social»,

---

16 Esto da lugar a lo que Tristan Harris ha llamado *human downgrading* (Thompson, 2019): el resultado de la interconexión de sistemas que producen daños que se refuerzan entre sí derivados del uso creciente de estas nuevas tecnologías de captado de la atención y que conduce a la degradación de nuestra capacidad de atención, salud mental, la calidad de nuestras creencias, nuestra tolerancia y civilidad, al aislamiento, la polarización, adicción, etc. No se trata de que haya unos vídeos polarizantes en Youtube o que algunos adolescentes pasen demasiadas horas en Tiktok y se resientan sus notas y/o su salud mental, sino, haciendo un análisis institucional y no meramente interactivo cómo el sistema se mueve de forma conjunta en una dirección preocupante.

Nicholas Carr hablaba (2008), años antes, de que Google nos está haciendo estúpidos. a partir del ejemplo de la creciente dificultad para concentrarse leyendo desde la aparición de internet, Google, las notificaciones, etc.

17 No es lo único, existe un creciente número de estudios sobre muchos aspectos que ya hoy se pueden predecir con un sorprendente grado de certeza, anticipar cuándo un trabajador va a dejar su trabajo (Rosenbaum, 2019), la tendencia al suicidio (Roy et al., 2020), cuándo una relación va a comenzar (Diuk, 2014), la tasa de divorcios o ruptura de parejas (Valenzuela, Harpern y Katz, 2014), la orientación sexual de las personas (Wang y Kosinski, 2018) o incluso la tendencia a trastornos de la alimentación (Sadeh-Sharvit, 2020).

18 Traducción propia.

19 Shriver, Nair y Hofstetter (2013, p. 27) «Mostramos que las políticas diseñadas para aumentar los lazos sociales pueden ayudar indirectamente a facilitar la generación del contenido, creando así un enlace entre la formación de lazos y los ingresos por anuncios». Traducción propia.

bombardeándonos con sugerencias de amistad sistemáticamente. Esta no es una característica necesaria de una red social, hay que programarla para que funcione y tiene que haber científicos, ingenieros y analistas de datos estudiando qué factores son los más relevantes a la hora de garantizar que una sugerencia será aceptada. Si se hace, es porque el límite que impone el número de Dunbar produciría una audiencia menor, lo que hace que nos sintamos menos motivados para generar contenido y, al haber menos contenido en la red social, pasaremos menos tiempo en ella. Así, nos encontramos con que la media de amistades por usuario en Facebook es de 338, número muy superior en los jóvenes, de los que aproximadamente una cuarta parte posee más de 500 amistades<sup>20</sup>. Las POBEA encuentran así modos de esquivar los límites de nuestros cerebros<sup>21</sup>.

Todo este modelo económico lo bautizó Zuboff (2020) como capitalismo de la vigilancia, un sistema en el que la principal fuente de riqueza y poder proviene de estos nuevos medios de acumulación de información<sup>22</sup> y modificación conductual, superando totalmente a la propiedad de los medios de producción (tierra, fábricas, capital, etc.) del capitalismo más tradicional. Zuboff, muy crítica con este modelo económico, considera que:

[...] el capitalismo de la vigilancia es una fuerza sin escrúpulos impulsada por unos novedosos imperativos económicos que ignoran las normas sociales y anulan los derechos elementales asociados a la autonomía individual y que tan imprescindibles resultan para que las sociedades democráticas siquiera sean posibles. (2020, 23)

Un requisito esencial de la economía de la atención es mantener a los usuarios dentro de las plataformas donde su atención genera datos, mejora los modelos y puede ser vendida a los anunciantes, por lo que estas tratan de optimizar sus sistemas de tal modo que los usuarios pasen la mayor cantidad de tiempo posible en ellas. El Senador B. Schatz de EE.UU. resumía algunos de los efectos sociales y políticos de las POBEA en una audiencia sobre el uso de los algoritmos y su influencia en el público:

Las redes sociales y otras plataformas ganan dinero manteniendo a los usuarios comprometidos (*engaged*), por lo que han contratado a los mejores ingenieros y mentes del mundo de la tecnología para conseguir que los usuarios permanezcan más tiempo dentro de sus aplicaciones y en sus plataformas *online*. Descubrieron que una forma de mantenernos a todos enganchados es utilizar algoritmos que nos alimenten

---

20 Datos de 2014, <https://www.pewresearch.org/fact-tank/2014/02/03/what-people-like-dislike-about-facebook>

21 Aunque podría resultar que solo lo consigan de forma superficial, favoreciendo la generación de escala, pero no el aumento de relaciones significativas reales. El propio Dunbar (2016, pp. 6-7) estudió la relación de las redes sociales con su hipótesis del cerebro social limitado, y sostiene que sus resultados parecen apuntar a que las redes sociales no aumentan el tamaño de las relaciones sociales que se pueden mantener de manera efectiva y que el análisis del tráfico de redes como Facebook o Twitter reproduce la estructura social de las relaciones sociales presenciales (*offline*).

22 El uso del término peyorativo «vigilancia» es intencionado. El registro sistemático, no siempre declarado, de datos de todas las métricas posibles convierte a los usuarios en sujetos vigilados. El enorme alcance al que llega la extracción de datos en un uso promedio de las nuevas tecnologías lo desglosa Véliz (2020) en el primer capítulo de su libro.

con un flujo constante de contenido cada vez más extremo e inflamatorio. Este contenido se inserta con muy poca transparencia o supervisión por parte de humanos. Esta configuración junto a aspectos básicos de la psicología humana nos hacen vulnerables a las mentiras, los engaños y la desinformación. Una investigación del Wall Street Journal el año pasado descubrió que el motor de recomendación de Youtube a menudo lleva a los usuarios a teorías de conspiración, puntos de vista partidistas y vídeos engañosos, incluso cuando los usuarios no buscan ese tipo de contenido y vimos que los algoritmos de YouTube recomendaban vídeos de niños después de que los usuarios vieran contenido sexualizado que no involucraba a niños. Este no es solo un problema de YouTube, vimos cómo todas las principales plataformas tenían problemas para contener la difusión de vídeos de la masacre de Christchurch y la propaganda anti-musulmana. El tiroteo se transmitió en directo en Facebook y se subieron más de 1 millón de copias en todas las plataformas. Muchas personas relataron haberlo visto en reproducción automática en sus feeds de redes sociales y no darse cuenta de lo que era [...]. Las empresas están dejando que los algoritmos actúen sin control.<sup>23</sup>

Es crucial insistir en que esta forma de explotación de la economía de la atención no es una tecnología concreta, no se reduce a sus algoritmos ni a los mecanismos que fomentan la adicción (que se expondrán más adelante, sección 4.2), sino a la lógica subyacente a estas tecnologías, que las convierte, desde una óptica mercantil, en las respuestas más racionales a los incentivos que ofrecen los mercados y las legislaciones actuales. Esta forma de la economía de la atención no puede darse fuera del mundo digital y de la acumulación masiva de datos, pero no es inherente al mundo digital en sí, que podría adoptar muchas otras formas diferentes. La configuración actual de las redes sociales y demás aplicaciones es tan solo una entre muchas, no la única versión posible de esas tecnologías. Podrán existir (existieron y existen) redes sociales con configuraciones diferentes si se transforman los incentivos que dan lugar a ellas y las lógicas legales, económicas y sociales en las que se desarrollan. La tecnología no es un ente en sí mismo que se desarrolla independientemente de la sociedad y la economía en la que se encuadra. No existe la inevitabilidad tecnológica, por lo que, si existiera una relación entre las POBEA con sus tecnologías predictivas y una serie de graves perjuicios para la humanidad, podrían y deberían modificarse las condiciones de base que dan lugar a esas tecnologías. En un contexto de economías de mercado, las empresas tenderán al desarrollo de tecnologías que respondan de maneras eficaces a los incentivos existentes y posibiliten la explotación eficaz de los recursos disponibles en su búsqueda de beneficios económicos. Uno de esos incentivos es precisamente la legislación (o la ausencia de la misma), por lo que crear legislación y modificar la que sea defectuosa puede contribuir a mitigar los efectos más negativos de la economía de la atención (como se propondrá en la sección 8.4).

---

<sup>23</sup> Comité de Comercio del Senado de los EE.UU.: <https://www.c-span.org/video/?462071-1/technology-companies-algorithms> Minuto 4:53 en adelante. Traducción propia.

#### 4. PROBLEMAS

La explotación de la atención humana genera o potencia una gran cantidad de problemas. Citemos unos pocos ejemplos antes de profundizar en las dos categorías fundamentales:

- Youtube y la promoción de la pedofilia (Fisher, 2019): vídeos familiares inocentes que incluían niñas pequeñas en bañador fueron recomendados sistemáticamente por el algoritmo de Youtube<sup>24</sup> a usuarios con tendencia a buscar vídeos de preadolescentes parcialmente vestidos, a menudo dirigiendo a los usuarios a estos vídeos familiares después de que vieran vídeos con contenido o temática sexual. Algunos pedófilos usaban la sección de comentarios para guiar las búsquedas de otros pedófilos, pero el algoritmo no llevaba únicamente a usuarios pedófilos a ese tipo de contenido, sino que otros usuarios que estuvieran viendo contenido erótico podían recibir recomendaciones de vídeos cuyos protagonistas eran cada vez más jóvenes, llevando incluso a sugerir vídeos de niñas entre 5 y 6 años en bañador, vistiéndose o abriéndose de piernas (como ejercicio de gimnasia acrobática). Youtube solo reaccionó cuando los anunciantes exigieron medidas, algo que enlaza con la discusión sobre la necesidad de limitar este asunto a una cuestión de responsabilidad social de las empresas, que trataremos más adelante.
- Facebook y el genocidio de Myanmar (Mozur, 2018): personal militar de Myanmar utilizó Facebook como herramienta para incitar a la limpieza étnica, principalmente de los Rohingya, a través de una campaña con cuentas falsas con las que inundaron las redes de contenidos que incitaban al asesinato y la violación. Estos contenidos podían ser tanto noticias falsas sobre ataques a mujeres y delitos comunes como llamadas directas a la acción violenta. Los oficiales encargados, que llegaron a ser hasta 700, se hacían pasar por cuentas de noticias o de personajes famosos y subían el contenido en los momentos de más impacto. El objetivo de la campaña era llevar a las comunidades a altos niveles de estrés social por miedo e ira que desembocasen en violencia. Facebook confirmó la existencia de esta campaña, que pasó desapercibida a sus filtros internos, y reconoció que habían reaccionado demasiado lentamente. Reclutar moderadores para estas redes sociales no resulta efectivo, ya que ni siquiera los 15.000 que tiene Facebook a sueldo son capaces de estar al día con la cantidad de contenido que se sube. Tampoco hablan los lenguajes de todos los territorios que utilizan Facebook en el mundo. Además, este trabajo de moderación causa profundos problemas psicológicos a algunos de estos trabajadores y la empresa ha tenido que compensarles económicamente por ello (Newton, 2019; Whittaker, 2020).
- Facebook y la relación con ataques a refugiados (Taub, 2018): el aumento de publicaciones sobre la crisis de los refugiados en las redes y el consiguiente aumento de ataques a estos llevó a

---

<sup>24</sup> Por ejemplo, el vídeo familiar de una niña y su amiga de 10 años bañándose en una piscina en el jardín alcanzó rápidamente las 400.000 visualizaciones, aunque la cuenta apenas tenía seguidores.



la realización de un estudio sobre la relación entre esas variables (Müller y Schwarz, 2018). El estudio comparó datos de riqueza, demografía, apoyo a la extrema derecha, venta de periódicos, número de refugiados, etc., hasta un total de 3000 variables y concluía que el único dato constante era que aquellas ciudades donde el uso de Facebook en particular (y no otros usos de internet) era mayor que la media experimentaban sistemáticamente más ataques que en el resto, por lo que determinaron que existía una correlación entre las redes sociales y los crímenes de odio, que aumentaban cuando subían las interacciones y caían cuando otros eventos informativos copaban las redes. No solo eso, sino que encontraron que las caídas de servicio de internet en áreas específicas se correlacionaban con caídas significativas de los ataques a refugiados, en la misma proporción que aumentaban los ataques con el uso intensivo de Facebook. Los algoritmos tienden a favorecer contenido extremo, ya que las emociones como el miedo y la ira son especialmente buenas para el objetivo maximizar el tiempo en la aplicación y la interacción<sup>25</sup>, algo que se conoce como extremismo algorítmico (*algorithm extremism*) y trataremos en la sección 5.4.

- Facebook segmenta anuncios de armas a los asaltantes al Capitolio: (Mac, 2021): Facebook colocó anuncios de armaduras corporales y otro equipamiento militar junto a contenidos conspirativos sobre las elecciones y el asalto al Capitolio, lo cual apunta a que Facebook se beneficia directamente del contenido que amplifica la polarización política estadounidense. Facebook no actuó, ni siquiera ante los avisos de sus propios empleados. Es particularmente preocupante que, a la vez que el algoritmo que selecciona contenidos alimenta los muros virtuales (*feeds*) de los usuarios más extremistas con desinformación y contenido conspiratorio, les muestra anuncios de equipación militar que apuntan específicamente a la posibilidad de conflictos civiles violentos. Formalmente, Facebook tiene la política de prohibir grupos y cuentas que promuevan abiertamente la guerra civil, pero muchos grupos consiguen burlar su vigilancia.
- El aumento de las muertes y accidentes de *daredevils* e imitadores (Elgan, 2019): Los *daredevils* son personas que se dedican a subir a sitios peligrosos<sup>26</sup>, a menudo prohibidos, para hacerse fotos y subirlas a redes en busca de viralidad, *likes*, seguidores y patrocinadores. Aunque la práctica de subirse a sitios peligrosos ha existido siempre, comenzó a hacerse especialmente

---

25 El contenido tóxico, polarizador, emocionalmente cargado nos atrae y nos incentiva a interactuar. La ira es la emoción que más nos mueve, además de que tiene el mayor potencial para generar reacciones en cadena. (Shaer, 2014). Pariser lo ilustra con un ejemplo (2011, p. 13): «Nuestros cuerpos están programados para consumir grasas y azúcares porque son escasos en la naturaleza... De la misma manera, estamos programados biológicamente para prestar atención a las cosas que estimulan: contenido que es grosero, violento o sexual y cotilleos que son humillantes, vergonzosos u ofensivos. Si no tenemos cuidado, desarrollaremos el equivalente psicológico de la obesidad.» Traducción propia.

26 También puede involucrar acciones arriesgadas con trenes, coches, motocicletas, animales salvajes o incluso la práctica del *extreme eating*, que tiene que ver con la ingesta masiva de comida, a veces en concursos, o de alimentos potencialmente peligrosos.

popular a partir de 2011, cuando las redes sociales convirtieron esta búsqueda de adrenalina en una fuente de fama e ingresos. Esto ha desembocado en un espectacular aumento de las muertes y las lesiones de estos *daredevils* y de sus imitadores, que no suelen tener la mismas habilidades y conocimientos<sup>27</sup>. Algunos especialistas entrevistados consideran que la relación con las redes sociales es evidente<sup>28</sup>. El asunto ha llegado tan lejos que en algunas zonas se han llevado a cabo campañas informativas sobre los peligros y se han colocado señales específicas de *no selfies*.

Ninguno de estos hechos responde a un intento consciente por parte de los gigantes tecnológicos. Ni Youtube busca el fomento de la pedofilia, ni Facebook quiere la limpieza étnica, las palizas a inmigrantes o la guerra civil, ni Instagram o Twitter pretenden que aumenten las muertes de los que buscan la fama *online*. Sin embargo, estos y otros resultados no son un accidente, sino la consecuencia necesaria de la estructura de esas redes sociales. Hay incluso quienes consideran (LaFrance, 2020), de forma hiperbólica, que la arquitectura actual de estas redes, especialmente Facebook, las convierte en Máquinas del Apocalipsis (*doomsday machines/devices*<sup>29</sup>). Para ella, la noción de que algunas redes sociales eran algo bueno en el pasado que se desviaron, pero que todo habría salido bien con unos pequeños ajustes (*fact-checking*, moderadores humanos, algo de regulación), es equivocada. Es la propia arquitectura dedicada a espolear ciertos comportamientos la que las convierte en dañinas. Las POBEA hacen exactamente lo que están diseñadas para hacer. No existen para buscar la verdad, generar un debate político cívico, defender los intereses de sus usuarios o someter a los poderes adecuada rendición de cuentas, si bien a veces pueden conseguir estas cosas como un subproducto. El surgimiento de fenómenos como QAnon son una conclusión lógica del diseño de estas plataformas, perfectas para amplificar y esparcir desinformación a gran velocidad. Su objetivo principal es la expansión del número de usuarios y del tiempo que pasan en su red y en ese sentido han tenido un éxito sin precedentes.

Los principales problemas que generan podemos asociarlos con la pérdida de la noción de la realidad. En un plano social, esto conduce al incremento de la posverdad y la polarización. En un plano individual, conduce a la adicción, el aislamiento, la depresión y la ansiedad.

No se trata de que estos problemas estén causados únicamente por las redes sociales. La polarización, la posverdad, las adicciones, los trastornos de la alimentación y demás problemáticas han existido antes del auge de la vida digital y la casuística particular de cada uno de ellos remite a múltiples causas. Sin embargo, las POBEA sí parecen realizar un papel facilitador o acelerador de estos fenómenos

---

27 Existe una lista en wikipedia con los accidentes. [https://en.wikipedia.org/wiki/List\\_of\\_selfie-related\\_injuries\\_and\\_deaths](https://en.wikipedia.org/wiki/List_of_selfie-related_injuries_and_deaths)

28 Ricky D. Crano, profesor de la Tufts University: «Estos incentivos para ganar en las redes sociales no son un accidente, sino que son producto del diseño de plataformas como Facebook, Instagram, and Twitter. Estos incentivos están en la propia receta y son esenciales para sus mecanismos corporativos de generación de beneficios» (Elgan, 2019). Traducción propia.

29 Dispositivo hipotético capaz de destruir toda la vida humana del planeta de forma automática y sin intervención humana (o incluso a pesar de los esfuerzos de esta para evitarlo). La idea de un dispositivo tal era hacer la guerra nuclear imposible de ganar y por lo tanto ineficaz, ya que habría una destrucción mutua asegurada.

#### 4.1 Irrealidad, posverdad y polarización

Se dice a menudo que vivimos en la posverdad, que las *fake news* dominan el panorama informativo de modo que cuesta distinguir lo que es cierto de lo que no. Esta situación se debe a la suma de muchos factores: el desarrollo de la economía de la atención, la proliferación de las redes sociales, el descrédito de los medios clásicos a medida que se iban convirtiendo en nichos ideológicos informativos o la propia caída del periodismo por la aparición de medios *online* gratuitos que hacen que la venta de prensa escrita o el periodismo de investigación en general resulten deficitarios frente a las plataformas de producción de contenidos por parte de usuarios.

Rápidamente se han sucedido las campañas de desinformación rusas con respecto a la invasión de Ucrania y el derribo del vuelo MH017 de Malaysia Airlines o la difusión masiva de la teoría conspiración QAnon y su principal exponente, el *Pizzagate*, que presuntamente revelaba la existencia de un *deep state* corrupto que compra y vende esclavos sexuales menores de edad y los distribuye desde el sótano de una pizzería en Washington DC. También son frecuentes las teorías al respecto del 5G, la relación entre las vacunas y el autismo o el negacionismo del Holocausto.

Sin embargo, algunos consideran que el panorama no es preocupante. Harari (2018) sostiene que la propaganda y la desinformación no son ninguna novedad. Naciones enteras han sido y siguen siendo negadas, como Palestina por Israel, Tíbet por China o en su momento toda la población aborigen de Australia por parte del Imperio Británico con la doctrina de *terra nullius*. «Si esta es la era de la posverdad, ¿cuándo tuvo lugar la era dorada de la verdad? [...] En realidad, los seres humanos siempre han vivido en la era de la posverdad.» (2018, p. 257)

A la propaganda anterior podríamos sumarle casos clásicos como la derrota de Upton Sinclair en California mediada por la difusión de informativos falsos rodados en el Hollywood de aquellos años 30 por los estudios MGM (Roland, 2016) o la influencia de Pulitzer y Hearst en la generación de la Guerra de Cuba (Hamilton, 2019). Nos podemos remontar más aún, a las historias de Hugh de Lincoln cayendo al pozo y siendo sacrificado por una camarilla de judíos y las persecuciones que generó, que desembocaron en última medida en su expulsión de muchos países. Las ficciones han acompañado al ser humano desde que hay memoria. De hecho, la visión antropológica de Harari se sustenta en que el dominio del ser humano sobre todas las demás especies se debe, fundamentalmente, a nuestra capacidad de crear ilusiones, creer en ellas y convencer al resto, incluso a extraños, de que hagan lo mismo para cooperar. El ser humano vive creyendo perpetuamente en ficciones como el dinero, las marcas, las religiones, el liberalismo, etc. Por lo tanto, crear y creer invenciones no es nada nuevo ni necesariamente negativo. Además, la capacidad de un grupo de creer relatos es una manera excelente de probar la lealtad de un grupo, de ver hasta qué punto colaborarán en los fines que el líder elija, frente a los hechos acreditados. Para un líder es mucho mejor como prueba pedirle a alguien que crea en un absurdo a que crea en la verdad. De hecho, hay una cierta relación de exclusión intrínseca entre la verdad y el poder.

La verdad y el poder pueden viajar juntos solo durante un trecho. Más tarde o más temprano, seguirán por sendas separadas. Si queremos poder, en algún momento tendremos que difundir ficciones. Si queremos saber la verdad sobre el mundo, en algún punto tendremos que renunciar al poder. (2018, p. 267)

Sin embargo, incluso alguien tan optimista como Harari traza el límite de los relatos en los que se puede creer. «En la práctica, el poder de la cooperación humana depende de un equilibrio delicado entre verdad y ficción. Si distorsionamos demasiado la realidad, nos debilitaremos, porque obraremos de maneras poco realistas.» (2018, p. 264)

Esta idea de que las *fake news* nos acompañan desde siempre no es exclusiva de Harari. También el historiador de la Antigua Roma, Francisco Pina, en su artículo (Segenni, S., 2019, pp. 74-91) sobre desinformación en tiempos de la República, defiende que la cuestión de las *fake news* no es propia de nuestra época, y tan solo ha variado el modo en que se difunden y llegan a la gente. Ahora vivimos en una época en la que cualquiera dispone de los medios para poner en circulación una noticia, sin la maquinaria que requiere una radio, un periódico o una televisión. Con estar conectados a internet, en cuestión de minutos podemos elaborar y compartir con infinidad de personas cualquier material. Muchos aprovechan esta capacidad para crear historias falsas. Esta desinformación, la difusión de información falsa o manipulada de forma deliberada con la intención de engañar a una audiencia para obtener beneficios personales, políticos o económicos suele tener la característica de presentar siempre contenidos sensacionalistas y desmesurados. Esta práctica actual ya ocurría también en la Roma clásica, ya que siempre ha habido gente dispuesta a creer en aquello que quería creer, siempre ha habido personas que ante cualquier noticia que sintonizase con sus sentimientos o ideología estaban dispuestos a creerla y repetirla. En Roma, donde no había medios de comunicación masivos y la sociedad tenía un acceso muy limitado a la educación, también se dieron fenómenos análogos a lo que hoy entendemos por *fake news*. Pina destaca tres: Primero, la conquista de Roma por César sin luchar, debido a los rumores que se propagaron según se acercaba. Segundo, el asesinato de Graco por la desinformación propagada por Quinto Pompeyo. Por último, la manipulación del testamento de Marco Antonio por Octaviano para convertirle en un traidor al servicio de Cleopatra. Destaca Pina, al igual que Harari, ese componente de irrealidad que hay en los bulos, mayor cuanto más lealtad se exige de un bando.

Como dice Casio Dión, el ambiente de indignación contra Antonio hacía creíble en Roma cualquier rumor, por implausible que éste fuera, como por ejemplo que Antonio deseaba trasladar la capital del Imperio a Alejandría, dando a Roma un papel secundario, algo que resultaba inaceptable para los romanos. (Segenni, 2019, p. 87)

Esta cuestión, en filosofía, nos acompaña desde prácticamente los orígenes, con las polémicas entre Sócrates o Platón y los sofistas, que, según aquellos, manipulaban mediante la retórica a sus audiencias. La sofística no podía conducir a la *sophía*, la verdadera sabiduría. Era por esa incapacidad de la gente

normal a alcanzar la verdad y distinguirla de las apariencias por lo que Platón no apoyaba la democracia. Sin embargo, algunos consideran que las críticas de Platón eran injustas, ya que los sofistas, en buena medida contribuyeron al funcionamiento de la democracia. Al fin y al cabo, su modelo se basaba en la búsqueda del consenso a través de la argumentación pública en el ágora. Si alguien quería triunfar, debía persuadir a la audiencia para creer en su forma de ver las cosas. En un artículo sobre la propaganda y la desinformación en los antiguos griegos, Rangappa y Mercieca (2020) reiteran esta idea de la irracionalidad frente a la lealtad:

Cuando la toma de decisiones se basa en "elegir un bando" en lugar de argumentaciones razonadas y el descubrimiento de la verdad, estas reglas básicas se anulan. Las investigaciones muestran que las creencias basadas en la lealtad, por ejemplo, a una persona o afiliaciones partidistas, son especialmente inmunes a los hechos que las cuestionan o refutan.<sup>30</sup>

Por lo tanto, quizá no haya una diferencia fundamental entre nuestros tiempos y los tiempos pasados que autorice a hablar de un tiempo nuevo de posverdad. Si acaso, habría una continuidad antropológica de tendencia a la desinformación.

Sin embargo, muchos otros autores discuten esto. Periodistas como Guillem Martínez (2018) o Daniel Gascón (2018) no están de acuerdo y creen que en este fenómeno clásico se ha producido un salto cualitativo y no solo cuantitativo. Martínez considera que la posverdad es la expresión de un espíritu de la época, una mecánica consistente en mezclar hechos reales con motivos sentimentales basado, principalmente, en la creación de identidades derivada de las *identity politics*, que nacieron con las izquierdas pero han sido recogidas y desarrolladas por el bando conservador. El componente fundamental que destaca Martínez no es el de la mentira, sino el de la sentimentalización<sup>31</sup>. Gascón sostiene también que la posverdad no es la mentira de siempre, sino una situación en la que los hechos objetivos carecen de importancia frente a la emoción o creencias personales. Sin embargo, a diferencia de escenarios donde la mentira se utilizaba para imponerse como verdad alternativa, en la situación actual se trata más bien de enfangar, de mezclar, de dificultar la percepción de diferencias, de cara a pavimentar un camino en el que la verdad sea tan solo lo que cada cual desee como verdad, lo que le oponga menos resistencias. Al igual que Martínez, destaca el componente identitario, que descansa en la suposición de que existe una experiencia últimamente incommunicable, cosa que se refuerza con los medios de expresión modernos, como Twitter o Facebook, que plantean un escenario de igualdad epistemológica, donde la viralidad es más importante que la verdad y la argumentación. Si el medio es

---

30 Traducción propia.

31 Presente incluso en las propias declaraciones de los ideólogos políticos actuales. Iván Redondo, ex-asesor del Partido Popular y principal asesor del Presidente del Gobierno Pedro Sánchez hasta julio de 2021, hizo estas afirmaciones: «Las tres emociones principales, con las que se puede jugar en campaña además, serían el miedo, el rechazo y la esperanza». «Estamos en un cambio de época. Ahora habría que decir “Son las emociones, estúpido”». «Entre otras cuestiones, yo antes me emociono y luego pienso. Primero siento y luego decido. Y esta es la manera de afrontar cualquier campaña electoral y cualquier cuestión en política» (recogido en Lenore, 2020).

el mensaje, las redes sociales emiten anti-intelectualismo. Tampoco son pocos los filósofos que relacionan la novedad de la posverdad con la herencia intelectual de la posmodernidad, como Kenan Malik (2017) o Lee McIntyre, que dedica a esta cuestión todo un capítulo de su obra *Post-Truth* (2018). En él, se estudia cómo la deconstrucción completa de todo discurso ha conducido a la apropiación por parte de aquellos que se sentían oprimidos por discursos hegemónicos, pero esto no ha ocurrido solo con los que realmente estaban oprimidos, sino también con otros que fingen opresión y se niegan a la discusión teórica<sup>32</sup>. McIntyre lo afronta desde el ejemplo de las empresas del tabaco pagando estudios artificiales para dudar del cáncer, las petroleras haciendo lo propio con el cambio climático o, en filosofía, la sociología de la ciencia de Latour, estudiando la crítica que supuso el experimento de Sokal. De la misma opinión es Robert Pennock (2010) que muestra cómo esto puede ser perfectamente utilizado para defender el creacionismo, el terraplanismo o lo que deseemos.

[...] Esto es como decir, no creas que la ciencia tiene nada que ver con la realidad; la evolución es solo una historia imaginativa. Resulta, sencillamente, que es la que cuenta la tribu científica. Desde la postura posmoderna radical, la ciencia no tiene ningún privilegio especial sobre ninguna otra visión acerca del mundo, incluso en lo relativo a cuestiones de hechos empíricos; cada tribu puede tomar su propia historia como punto de partida para sus demás creencias. Los creacionistas del Diseño Inteligente están igualmente legitimados cuando toman la creación divina y su voluntad sobre el hombre como su suposición inicial. (2010, p. 762)<sup>33</sup>

Así, cualquier experto que pretenda «saber» algo, no es más que alguien que quiere imponer ilegítimamente una postura. Lo que nos queda es creer, apostar, sentir, pero no atender a razones, porque presuntamente el que busca convencer busca siempre someter.

Esto significa que, a la base de todas las afirmaciones de conocimiento solo hay realmente un ejercicio de autoridad; son una táctica de intimidación utilizada por los poderosos para obligar a los más débiles a aceptar sus puntos de vista ideológicos. Dado que no existe la «verdad», cualquiera que diga «saber» algo en realidad está tratando de oprimirnos, no de educarnos. Tener poder nos permite controlar lo que es

---

32 Almagro, Osorio y Villanueva (2021, p. 45) llaman a esto apropiación epistémica a través del desvío epistémico. Ocurre cuando discursos desarrollados en los márgenes de la sociedad son utilizados por los relatos dominantes de los poderosos. Estas técnicas retóricas se utilizan para fomentar los desacuerdos cruzados y están relacionadas con el aumento de la polarización.

También trata Pina la relación entre las emociones y los bulos, especialmente al servicio de los poderosos: «Hay por lo general una estrecha relación entre bulo, rumor y miedo, miedo sobre todo a perder lo que se tiene, sea una casa, sea un privilegio. Adicionalmente, el miedo suele desembocar en enfado – incluso odio – contra los individuos que, presuntamente, ponen en riesgo los bienes o prebendas de los ciudadanos. La indignación, a su vez, genera acción: el enfado activa el deseo de castigar a quien ha sido identificado como enemigo, una vez interiorizado el peligro y el consiguiente temor. El bulo entendido como noticia está en el origen del rumor que permite modelar la opinión pública y contagiar el pánico, a partir del cual era factible justificar la muerte de Graco, la represión de los catilenarios o la guerra (civil) contra Antonio.» Segenni, (2019, p. 88)

33 Traducción propia.

verdad, no al revés. Si hay muchas perspectivas, insistir en que aceptemos alguna en particular es una forma de fascismo. (McIntyre, 2018, cap. 6)<sup>34</sup>

McIntyre destaca que lo nuevo de la posverdad, frente a la mentira de toda la vida, es que lo que está ocurriendo no supone un reto a una idea concreta de realidad o la capacidad de algunos de conocer la realidad, sino a la propia idea de la existencia de una realidad objetiva en sí misma. La realidad está esencialmente politizada, ideologizada y no cabe razonamiento sobre ella. Autores como Jason Stanley (2020) también advierten de este movimiento y señalan que son precisamente la irrealidad y el anti-intelectualismo dos de las diez estrategias fundamentales del fascismo. La posverdad es el fin de los paradigmas de la Ilustración, la renuncia a la conquista de lo universal por la racionalidad humana, de modo que solo quedan las perspectivas de cada cual. Sin embargo, dado que algunas percepciones de la realidad implican consecuencias indeseables (algo que incluso los optimistas como Harari mantienen<sup>35</sup>), parece una senda crecientemente peligrosa por la que adentrarse. Advertía Arendt (1998):

El sujeto ideal de la dominación totalitaria no es el nazi convencido o el comunista convencido, sino las personas para quienes ya no existen la distinción entre el hecho y la ficción (es decir, la realidad empírica) y la distinción entre lo verdadero y lo falso (es decir, las normas del pensamiento). (1998, p. 379)

La red funciona como ese espacio en el que hemos desterrado todo concepto de verdad, en el que cada cual tiene «su verdad» y es válida porque la dice, donde no se piden cuentas sobre los hechos, sino que las opiniones se validan y fundamentan principalmente en base a sentimientos, lealtades y sesgos previos. Esto supone un riesgo social que nos acerca hacia modelos autoritarios. Snyder (2017) va más allá al decir lo siguiente: «La posverdad es el prefascismo.» (2017, cap. 10). Como decía Mitch McConnell, la convivencia democrática requiere un compromiso colectivo con la verdad<sup>36</sup>. La existencia de dos bandos que habitan realidades cognitivas diferentes solo potencia la hostilidad. Esta polarización degrada el debate ciudadano hasta el punto de que hace imposible la comunicación. La agente de inteligencia Yaël Eisenstat (2016) denunciaba que le resultaba más sencillo sentarse con enemigos jurados de EEUU a discutir sobre política exterior que debatir con su vecino anti-abortista. Llamativamente, establecer diálogo con personas que tienen todo un mundo de información a su

---

34 Traducción propia. Entiéndase que McIntyre está criticando esta forma de entender la verdad.

35 Harari, Y. N. (2018, p. 268) «Todo esto no significa que las noticias falsas no sean un problema grave, o que políticos y sacerdotes tengan licencia para mentir con malicia. También sería completamente erróneo llegar a la conclusión de que todo son noticias falsas, que cualquier intento de descubrir la verdad está condenado al fracaso y que no hay ningún tipo de diferencia entre el periodismo serio y la propaganda. A todas las noticias falsas subyacen hechos reales y sufrimiento real. En Ucrania, por ejemplo, los soldados rusos están luchando de verdad, miles de personas han muerto de verdad y cientos de miles han perdido de verdad su casa. El sufrimiento humano suele generarse por creer en la ficción, pero el propio sufrimiento sigue siendo real.»

36 Si creemos en la democracia es porque entendemos que tiene sentido otorgar el poder a la gente que cumple ciertas condiciones de información y racionalidad básicas (por eso se le niega a los dementes y a los niños, por ejemplo) y de libertad (el voto es secreto y existen protecciones frente a la coacción). Precisamente Platón consideraba que la democracia carecía de valor porque no creía que pudiera funcionar un sistema en el que ciudadanos incapaces de reconocer la verdad deben ser los que articulen los mecanismos de una polis justa.

disposición y eligen no usarlo es más difícil que con aquellos que tienen un acceso muy limitado a ella y habían construido sus visiones del mundo a partir de ahí.

A la horizontalidad epistemológica, la velocidad a la que se pueden crear y transmitir bulos y la creación de filtros burbuja (sobre lo que se profundizará en la sección 4.1.1) hay que añadir un cambio notable en la forma en la que nos relacionamos con la información. En parte, esto también se debe a las fuentes de las que la gente extrae esa información con la que construyen su modelo del mundo. De acuerdo con un estudio en EE.UU. (Matsa y Shearer, 2018), el 68% de la población elige las redes sociales como fuente de información y casi el 25% lo hace como fuente primaria. Además, de forma especialmente llamativa, entre los que prefieren las redes sociales como forma de información, existe una confianza mucho mayor en ese medio que en otros, como la televisión, la radio o los periódicos. En estas redes, los usuarios actúan como receptores y emisores de información, a menudo no verificada, que puede viralizarse compartiéndose a velocidades impensables hace tan solo una década. La confianza en la información de internet y de las redes sociales es inferior en la Unión Europea, ya que acuerdo con datos del Eurobarómetro (2017), la confianza oscila entre el 20% y el 30%, mientras que en medios como la radio o la televisión supera el 50%, si bien los datos varían enormemente por países. Además, aproximadamente el 20% de la población usa las redes sociales como fuente primaria de información, e internet en general sube hasta el 42%, siendo la segunda fuente primaria, solo por detrás de la televisión. Más llamativos aún son los informes sobre la competencia ciudadana para distinguir *fake news* de noticias reales. En un estudio español con la colaboración de la Universidad Complutense de Madrid (Lógica, S., y UCM, 2017) se mostró que aunque el 60% se siente capacitado para distinguir adecuadamente las *fake news* de las noticias reales, tan solo el 14% de los individuos lo consigue.

Este contexto de posverdad aumenta la polarización. Este es un fenómeno antiguo que remite al tribalismo pre-histórico, por lo que no se trata de sugerir que haya surgido con las POBEA, pero sí se puede exacerbar o mitigar en función de las regulaciones y responsabilidades que recaigan sobre los medios y la arquitectura de las redes sociales que utilicemos<sup>37</sup>. Uno de los principales mecanismos que favorece la polarización es la creación de los filtros burbujas (*filter bubbles*).

#### 4.1.1 Filtro burbuja

«Filtro burbuja» es un término acuñado por Eli Pariser (2011) para explicar un fenómeno que surgió en 2009, cuando Google cambió su motor de búsqueda de tal manera que usase diferentes datos (en principio eran hasta 57 puntos diferentes) para proporcionar resultados personalizados. De este modo,

---

<sup>37</sup> Por ejemplo, en EEUU existía la *Fairness Doctrine* (1949-1987) de la FCC, que obligaba a las cadenas a mostrar ambos lados de la discusión en todo debate político y cuya retirada dio paso a la proliferación de medios de comunicación ultra-partidistas como Fox News y otros que han surgido en su estela. En relación con las redes sociales, podemos fijarnos en el caso de Infowars, un medio principalmente de *fake news* dirigido por Alex Jones, que, aunque relativamente pequeño, fue desproporcionadamente beneficiado por el algoritmo de recomendaciones de Youtube, ya que sus vídeos eran excelentes para aumentar el tiempo de visualización y el compromiso de los usuarios.



dos personas que busquen información sobre un suceso serán remitidas a diferentes sitios web. Estos motores personalizados han sido replicados por Facebook, Instagram, Youtube, Netflix, Amazon, etc<sup>38</sup>. Su éxito radica en que son muy útiles para estructurar la sobreabundancia de información en la que nos movemos<sup>39</sup>. Nuestra atención y nuestro tiempo no bastan para procesarlo todo, de tal manera que estos algoritmos nos ayudan a encontrar lo más relevante para nuestras vidas, gustos, búsquedas, etc., ajustando la catarata informativa a la medida de cada uno. Esto provoca una percepción diferente de la realidad que nos rodea, tanto más cuanto más nos movamos en el mundo digital. Bastaría con imaginar una enciclopedia que nos diese artículos personalizados (y con información contradictoria entre sí) en función de lo que su algoritmo estimase que deseamos recibir. De esta forma, los filtros burbuja son los universos digitales únicos de cada cual, ajustados en función de su comportamiento *online*. El problema es que no decidimos qué atraviesa el filtro ni sabemos qué se queda fuera, por lo que nos sitúan en estados de aislamiento intelectual privados<sup>40</sup>. Esto afecta tanto a la forma en la que recibimos información como a nuestra manera de procesarla e incluso a cómo pensamos: «Como una lente, el filtro burbuja transforma de manera invisible el mundo que experimentamos al controlar lo que vemos y lo que no vemos. Interfiere en la interacción entre nuestros procesos mentales y nuestro entorno externo.» Parisier (2011, p. 48)<sup>41</sup>. Los filtros burbuja operan por medio de un doble proceso. Primero, nos rodean de ideas con las que estamos familiarizados o nos sentimos cómodos, creando una confianza excesiva en nuestros marcos cognitivos. Segundo, eliminan de nuestra vista algunos de los principales estímulos que nos llevan a querer aprender: las ideas que nos confunden, nos cuestionan, nos incomodan.

Esta hipótesis está siendo actualmente discutida. Cardenal et al. (2019), estudiando el caso español durante un periodo de 3 meses en 2015, encontraron que la tendencia de los usuarios es la de consumir más información en medios de su tendencia ideológica, pero relacionándose también de manera significativa con medios del otro lado del espectro, especialmente para aquellos que se identifican con la izquierda, por lo que consideran que la hipótesis de las cámaras de eco partidistas es un mito (2019, p. 13), si bien destacan que su estudio sugiere que el comportamiento polarizado aumenta cuando aumenta

---

38 Facebook, por ejemplo, solía mostrar en el *feed* el contenido de los amigos y grupos a los que se pertenecía por un orden estrictamente cronológico, lo último subido por cualquiera de ellos era lo primero que se veía. Fue sustituido por un sistema que trataba de mostrar los contenidos que tuvieran más posibilidad de interacción, de tal modo que, en la práctica, las publicaciones de muchos contactos desaparecían completamente de los muros.

39 Parisier (2011, p. 11) comenta que si sumásemos toda la información comunicativa que ha generado la humanidad desde sus comienzos hasta 2003, tendríamos la misma cantidad de información que a día de hoy (fecha de la publicación del libro) se genera cada dos días.

40 En cierta medida, el filtro burbuja es un tipo de cámara de eco (*echo chamber*) digital, que, a diferencia de las cámaras de eco tradicionales en las que tienden a moverse los seres humanos, está inscrita a priori en la estructura de la POBEA y sin opción para el usuario de salirse de ella. Un ejemplo de cámara de eco tradicional sería la que crean los individuos cuando eligen ver ciertos canales de televisión, leer las noticias en ciertos periódicos y relacionarse con amigos de ideas afines. Sin embargo, a diferencia de los filtros burbuja, los individuos conocen la existencia de otros canales, periódicos y personas, que a menudo ven y escuchan, aunque no les guste. El filtro burbuja directamente hace un corte de la realidad, de tal manera que el usuario ni siquiera puede elegir ignorar tal información. De esta manera, elimina la agencia individual de la ecuación en busca de la maximización de la atención.

41 Traducción propia.

el consumo de noticias *online*<sup>42</sup>. Una línea similar adoptan Young y Miller (2021) al estudiar el mundo anglosajón, concluyendo que las cámaras de eco no son generalizadas, aunque sí existen algunos nichos de radicalización.

Además del empobrecimiento intelectual, los filtros burbuja pueden llevar a la polarización y, posteriormente, a la violencia. Las personas tienden a amoldarse a las normas sociales de manera intuitiva, lo que supone un freno a la posible conducta antisocial. Para ello, perciben las creencias de la gente de su alrededor y adecuan o moderan sus propios pensamientos. La generación de filtros burbuja destruye este proceso aislando a los individuos de las voces moderadoras y sumergiéndolos en grupos de ideas afines, de tal modo que pueden llegar a la conclusión de que mucha más gente comparte sus ideas. En el caso de ideas hostiles, un individuo puede llegar a creer que existe un apoyo generalizado a su rechazo hacia otros individuos y darle la impresión de que, si actuase, tendría un fuerte apoyo social, precipitando así acciones violentas. Los filtros burbuja hacen incluso difícil detectar la propaganda política violenta. En otras épocas era imposible segmentar los mensajes con la precisión actual, pero ahora podemos desconocer completamente las experiencias personalizadas de nuestros vecinos.

#### 4.1.2 El medio es el mensaje

La hipótesis Sapir-Whorf sostiene que las características gramaticales y léxicas de un idioma determinan o, al menos, influyen en la visión del mundo que posee una comunidad lingüística. En su versión fuerte (determinación), la hipótesis está muy desacreditada, pero hay cierto consenso (Chandler, 1995, p. 18) en que opera en su forma débil (influencia). Análogamente, existe un lugar común en teoría de la comunicación que fue acuñado por Marshall McLuhan (1994): «el medio es el mensaje». Esta idea sostiene que medio y mensaje funcionan como una pareja, que el medio influye en el mensaje de manera que los códigos comunicativos que funcionan en una novela no se corresponden necesariamente con los que funcionan en una película en el cine o en un *podcast* de radio *online*. El canal por el que discurre un mensaje tiene un peso fundamental en el contenido y la recepción del mismo, igual que el idioma tiene una cierta influencia en la estructuración cognitiva que hacen del mundo sus hablantes.

Desde el estudio de las redes sociales es fundamental tener esto en cuenta, ya que los mecanismos epistemológicos que operan en los entornos académicos o en el mundo del periodismo tradicional se ven transformados. En su análisis de la influencia de ciertos *hashtags* de Twitter en la construcción de una realidad alternativa, Martín (2021) concluye que en la forma concreta de difusión que opera en esta red social «el *hashtag* es el mensaje». Los principales prescriptores aglutinan en torno a estas etiquetas

---

42 ¿Quizá esto apunta a que quienes tienen mejor acostumbrado el algoritmo a sus gustos tienen a polarizarse más? ¿No apoya esto la teoría de los filtros burbuja? En cualquier caso, cuanto más tienda un sistema a guiar a los usuarios hacia contenidos en los que no sientan fricciones, más posible será la creación de filtros burbuja. La hipótesis de los filtros burbuja funciona solo para el acceso a medios que se realiza desde POBEA, pero si los usuarios introducen manualmente las URLs de los medios a los que quieren acceder, esquivan el problema. Sea como fuere, se podría propiciar un diseño de las POBEA que fomente el salto entre medios de ideologías políticas no afines.

mensajes que tienen gran alcance porque sustituyen las formas de verificación tradicionales con una auténtica cámara de eco<sup>43</sup>. El *hashtag* se transforma en la línea guía que señala el camino hasta ese espacio en el que los mismos mensajes se van a repetir una y otra vez, con disonancias mínimas. En otro tiempo, un ciudadano podía rodearse de los periódicos y radios que desease, pero tenía que vivir en comunidades excepcionalmente homogéneas para no tener que escuchar otro tipo de informaciones en su día a día. Hoy, puede incluso participar en el discurso global y la resonancia de su cámara ya que los canales que ofrecen las redes sociales son multilaterales. Un usuario lee el *hashtag*, escribe su propio mensaje, retuitea otro y le da *like* a varios y, de esta forma, los receptores se vuelven a su vez emisores, reforzando su confianza en el mensaje difundido a través del *hashtag* por medio del compromiso adquirido con su participación. Estos mensajes, además, se fundamentan en argumentaciones emotivas en los que la validación no requiere de los hechos, sino de una supuesta percepción de los mismos. De esta forma, como destaca el artículo, las herramientas de las redes sociales pueden crear entornos en los que millones de votantes estén convencidos de que ha habido un fraude electoral masivo en unas elecciones en las que su candidato, que habría obtenido una mayoría abrumadora, ha sido sustituido por un plan de las élites globales y la única alternativa democráticamente responsable que queda es el asalto a las instituciones. También pueden crearse entornos en los que se afirme rutinariamente que la Tierra es plana, Finlandia no existe, las vacunas son malas o existe un genocidio sistemático contra los blancos en Occidente.

Lo llamativo no es que los seres humanos puedan creer ficciones totalmente separadas de la realidad, esto ha ocurrido en todas las épocas y, a menudo, con consecuencias históricamente catastróficas. Sin embargo, en otros tiempos, los que llevaban a cabo el control de las narrativas tenían que desplegar un dominio casi absoluto de los discursos generales de la sociedad para poder persuadir de su mensaje a la mayoría. Censura, propaganda, creación de medios afines, represión, violencia, etc., se volvían imprescindibles para la generación de cámaras de eco que permitiesen el gobierno de los discursos. Ahora, en cambio, la propia estructura de las redes sociales permite lo mismo en un entorno de pluralidad mediática. El usuario de Twitter tiene acceso teórico a informaciones que situarían a un terraplanista ante potentes disonancias cognitivas, pero si vehicula su camino en la red a través de los *hashtags* adecuados propagados por los principales prescriptores de contenidos que sigue y la red le sugiere, no tiene por qué enfrentarse jamás a ello. Además, si el objetivo de una POBEA es mantener la atención del usuario ante la pantalla, deberá garantizar su confort epistemológico dentro de su cámara de eco, y no existe ningún incentivo para que la red social proporcione visiones alternativas, sino, al contrario, la consecuencia lógica es que se elaboren algoritmos que potencien la creación de nichos para maximizar los beneficios económicos derivados de ese modelo de negocio. En ese sentido, la estructura de POBEA como Facebook o Twitter generan más polarización en el debate intelectual, político o social

---

43 «La *Alt-Right* muestra abiertamente su falta de confianza hacia los medios de comunicación, centros de investigación, expertos, etc. Para compensar esta inseguridad informativa, ofrece su propia información, transmitida a través de medios y figuras afines.» (Martín, 2021)

que otras como Kialo, cuyos mecanismos son muy diferentes<sup>44</sup>.

Estos son algunos de los riesgos sociales que genera la economía de la atención<sup>45</sup>. Ahora se presentarán algunos de los individuales.

## 4.2 Irrealidad, adicción y trastornos

Los principales problemas individuales tienen que ver con la adicción. Ya se ha visto que incentivar la adicción es una consecuencia natural de la economía de la atención. En este apartado se desarrollarán dos aspectos: primero, la relación entre la adicción y trastornos psicológicos importantes, como la depresión o la ansiedad; y segundo, algunos mecanismos mediante los que se fomenta la adicción.

Lo primero que hay que abordar es si nos encontramos ante un problema. Históricamente ha habido toda una serie de pánicos morales en los que se culpaba a diferentes elementos culturales nuevos de muchos males sociales: los cómics, las películas, los videojuegos o incluso las novelas (North, 2014). Por lo tanto, no basta con señalar el simultáneo aumento de las tasas de algunos trastornos psicológicos y la aparición de las redes sociales, sino que hay que profundizar más. Lo hace Twenge (2017, caps. 3 y 4) centrándose en las nuevas generaciones que se han criado en entornos de normalización de las redes sociales mostrando muchas correlaciones relevantes<sup>46</sup> en las que los datos parecen apuntar siempre claramente a las redes sociales como el mayor factor de riesgo. Estos datos han sido discutidos (Orben y Przybylski, 2019)<sup>47</sup>, pero se ha criticado que en su análisis no distinguen los usos del tiempo de pantalla<sup>48</sup> que miden, ni separan por géneros. A lo que apuntan respuestas a su estudio (Haidt y Allen,

---

44 Kialo está articulada como un sistema de afirmaciones y argumentos que las apoyan o las critican (que, a su vez, pueden cimentarse en otros argumentos que los apoyen o critiquen, encadenándose así hasta el nivel de profundidad que los usuarios sean capaces de alcanzar colectivamente). Las afirmaciones o los argumentos no vienen apoyados por la autoridad de los usuarios, aunque, como en Wikipedia, se puede acceder al historial de edición y la discusión que hay detrás de cada uno, donde sí consta la autoría. Se puede observar el desarrollo pacífico de un tema tan emocionalmente complicado (y a menudo violento) como es la existencia de Dios en <https://www.kialo.com/the-existence-of-god-2629>. Además, Kialo no posee ninguna de las características que se mencionarán en el apartado 4.2.

45 Hay muchos otros. Sin entrar a desgranarlos en profundidad, merece la pena mencionar dos que tienen la capacidad de empeorar el panorama con respecto a la irrealidad, la posverdad y la polarización drásticamente, especialmente en un contexto de ausencia de regulación y responsabilidad de las redes sociales: Por una parte, la mejora de la tecnología para la producción de *deepfakes* permite ya la producción de vídeos y audios falsificados difícilmente distinguibles de otros auténticos, y la tendencia es que la mejora hará que a muy corto plazo se vuelva accesible a bajo presupuesto y cantidad de trabajo. Por otra, la multiplicación de los sesgos que adquieren los algoritmos cuando se los alimenta con información defectuosa o sesgada. Este problema tendrá tanta más importancia cuanto más se generalice el uso de algoritmos para funciones sociales relevantes, como la justicia o la medicina. Un ejemplo de la influencia en la justicia está en Angwin et al. (2016) y en medicina en Ledford (2019). Con respecto a los *deepfakes*, aún en pañales, podemos ver un ejemplo aquí <https://www.youtube.com/watch?v=cQ54GDm1eL0> producido con muy pocos medios y otro mucho más sofisticado aquí <https://www.youtube.com/watch?v=Yewm6TfLZ3Q> con intenciones comerciales.

46 Por ejemplo: caída en el porcentaje de adolescentes que van a una fiesta o más al mes, el porcentaje de adolescentes que queda con amigos a diario o el porcentaje que acude a un centro comercial al menos una vez al mes; aumento del porcentaje de adolescentes que pasa más de 10 horas a la semana *online*, mayor factor de riesgo relativo de infelicidad, soledad, depresión y suicidio comparando el uso de redes sociales con otras actividades y en función del tiempo de uso, etc.

47 En los que, utilizando una enorme base de datos, muestran que el tiempo de pantalla influye lo mismo que comer patatas o llevar gafas.

48 Toman *screen time* como genérico, sin distinguir entre ver vídeos, jugar a videojuegos, redes sociales, etc.

2020; Twenge et al., 2020) es a que el factor relevante es específicamente el tiempo en redes sociales, indicando además que el crecimiento de muchas de estas tasas no es lineal, sino exponencial, lo que significa que los riesgos están principalmente relacionados con un uso intensivo de las redes sociales y no con uno esporádico, que podría ser sano y pro-social. Twenge (2017, cap. 3) indica que los riesgos comienzan con usos superiores a 2 horas y que usos de 5 o más están claramente unidos a altos riesgos de infelicidad y suicidio. Además, los problemas de ansiedad y depresión derivados de este uso afectan desproporcionadamente a chicas, adolescentes y mujeres jóvenes<sup>49</sup>. Se podría argumentar que más que una crisis de salud mental, estamos simplemente ante una generación que tiene más información y por lo tanto pide ayuda antes, con lo que muchos trastornos que han existido siempre se diagnostican antes y en mayor cantidad. Sin embargo, el hecho de que no solo crezcan los diagnósticos, sino también las hospitalizaciones y los suicidios<sup>50</sup> apunta en la dirección de una crisis. Las redes convierten el acoso escolar en cyberbullying que se extiende a todo el día. El FOMO<sup>51</sup> que genera ver las vidas aparentemente plenas y perfectas de los demás es constante, aumentando la envidia, el resentimiento y la insatisfacción (Krasnova et al., 2013), exacerbado porque la comparación social se hace con personas cercanas y parecidas. El número de adolescentes que contestan que no pueden hacer nada bien, que su vida no es útil o que no la disfrutan, que se sienten inferiores a otros, que sus amigos tienen vidas mejores crece drásticamente a partir de 2011 (Twenge, 2017, cap. 4), al hilo del crecimiento de las tasas de depresión. También es drástico el descenso en las horas de sueño nocturno. Algunas personas han relacionado los aumentos en las tasas de depresión y ansiedad con la presión escolar, pero los datos parecen apuntar a que precisamente los estudiantes que pasan más tiempo haciendo deberes son menos susceptibles de estos problemas.

Por último, hay que mencionar algunos de los principales mecanismos que fomentan las adicciones del comportamiento, para los que usaremos lo aportado por Schüll (2012), Eyal (2014) y Alter (2018):

- Pequeños objetivos: Las tareas grandes tienden a hacer sentir impotencia y lejanía. Poblar una aplicación de micro-objetivos fáciles de lograr, uno tras otro, aumenta la dedicación de los usuarios. La obsesión por el cumplimiento de objetivos está fuertemente asociada con la infelicidad, y muchos expertos consideran que las POBEA los promueven de forma sistemática (Alter 2018, capítulo 4)<sup>52</sup>. Los micro-objetivos pueden ir subiendo progresivamente de

---

49 La presión social por el aspecto físico parece ser un factor determinante y afecta mucho más a las mujeres que a los hombres. Es habitual que muchas chicas jóvenes aprendan pronto que reciben más *likes* y atención si suben fotos con más piel expuesta, también que eliminen fotos si no reciben la cantidad esperada de *likes*.

En cualquier caso, otros estudios, como el de Kross et al. (2013), que también apoyan la tesis de que el uso de redes (Facebook específicamente) es un buen predictor de la infelicidad, no encuentran variación por género.

50 Estos datos también se aplican a España, donde el suicidio ya es la principal causa de muerte de los jóvenes (Paniagua, 2021). Además, aunque los datos todavía son provisionales, todo apunta a que la pandemia ha deteriorado la salud mental aún más, coincidiendo con el incremento de uso de las redes sociales desde el confinamiento (Juste, 2021).

51 *Fear of missing out*, miedo a estarse perdiendo algo (bueno, interesante, necesario).

52 Ejemplos: Contar pasos, calorías, puntos, *likes*, seguidores, logros desbloqueables, rachas, etc. Snapchat, por ejemplo, contabilizaba los días consecutivos compartiendo fotos. Algunas *apps* de ejercicio hacen lo propio y se ha informado de lesiones por estrés físico debido a que muchos usuarios querían continuar sus rachas.

dificultad, de tal manera que su facilidad inicial resulte un cebo atractivo, pero su dificultad posterior requiera de más y más tiempo, generando así un cierto espejismo de progreso que mantiene la necesidad de continuar.

- Retroalimentación rápida (*Fast feedback*): El *feedback* es esencial para el aprendizaje de los humanos y en general disfrutamos cuando observamos y comprendemos los resultados de nuestras acciones. La atracción por el *feedback* puede ser asombrosamente adictiva<sup>53</sup>. Los botones de «me gusta», por ejemplo, proporcionan *feedback* en tiempo real, igual que las métricas de uso de Youtube. También los pequeños sonidos, cambios de colores, las notificaciones, etc. En este sentido, fue clave la introducción de los *likes* y *follows* para la explosión de las redes sociales<sup>54</sup>, que aportaban validación social instantánea<sup>55</sup>. Estos, además, no solo generan atención, sino que generan la adicción a la atención, por lo que aumentan la adicción de los usuarios y también de los creadores de contenidos. La velocidad también es esencial para las adicciones<sup>56</sup>, y las POBEA están diseñadas para proporcionar *feedback* inmediato.
- Recompensas aleatorias. Experimentos con ratones y con palomas demostraron lo que luego se ha explotado en casinos y redes sociales: los individuos se obsesionan más con sistemas de recompensa aleatorios que lineales. En algunas aplicaciones, existe la función *pull-to-refresh*<sup>57</sup>,

---

53 Un ejemplo muy llamativo fue el experimento social del Botón de Reddit, que generó toda una subcultura y millones de personas unidas dedicando su tiempo para ver qué pasaba:  
[https://en.wikipedia.org/wiki/The\\_Button\\_\(Reddit\)](https://en.wikipedia.org/wiki/The_Button_(Reddit))

54 Casi a la vez, en 2009 introdujo Facebook el botón de *like*, que Twitter copió, y esta creó el de *retweet* que Facebook adaptó. Antes de eso, las redes sociales tenían la forma de Myspace, espacios en los que cada uno se exhibía públicamente con poca jerarquización social. No solo permitió tener un *feedback* de popularidad adictivo para los usuarios, sino que potenció la economía de la atención proporcionando enormes cantidades de información sobre aquello que interesaba a la gente, les hacía pasar tiempo en la *app* y aumentaba su compromiso.

55 Es muy llamativo cómo Tiktok ha arrasado entre los más jóvenes y muchos especulan que es por inflar artificialmente el *feedback* social, especialmente durante sus primeras publicaciones, para así engancharles con una mayor sensación de éxito y fama. Las métricas de impacto de la mayoría de usuarios y su número de seguidores, incluso en cuentas nuevas, son varias veces superiores a las que obtienen en otras redes como Instagram, incluso cuando ya tenían cuentas en ellas con anterioridad y por lo tanto una base de seguidores ya hecha, por lo que TikTok consigue generar artificialmente la sensación de éxito social que todos, pero especialmente los adolescentes, ansían. Esto lo hacen no especificando claramente el *feedback* recibido, no distinguen entre *views*, *likes*, etc., sino que ponen un corazón con un número al lado. De esta forma, la red social aprovecha las mecánicas adolescentes de necesidad de aprobación social, dando a sus usuarios un motivo para quedarse en la red social o incluso para sustituir sus otras redes sociales por esta. Su éxito, en este sentido, resulta especialmente preocupante, porque sería un estímulo para desatar una carrera entre plataformas hasta llegar a los límites del inflado artificial de la popularidad, generando cada vez más plataformas que, en su búsqueda de entregar a los usuarios jóvenes la validación social que buscan, creen recompensas sociales cada vez más alejadas de la realidad y que, sin embargo, por hacer sentirse bien a sus usuarios, las volverían más adictivas. Extraído de una charla entre J. Haidt, psicólogo social de la Universidad de Nueva York, y T. Harris, co-fundador del Center for Humane Technology y especialista en ética tecnológica: <https://www.humanetech.com/podcast/26-are-the-kids-alright>. En cualquier caso, no hay artículos académicos al respecto y no es probable que las redes sociales como TikTok liberen su algoritmo para comprobar ese posible *doping* social.

56 Por eso no existen comportamientos adictivos que se satisfagan en periodos de semanas. La adicción requiere de *feedback* inmediato para que el cerebro relacione el estímulo con el control del malestar que creaba la necesidad, lo que explica que la gente no pueda llegar a desarrollar un patrón patológico de visitas a museos, pero sí de consumo de cocaína.

57 Mecanismo por el cual al deslizar la pantalla hacia abajo vuelve a cargar los contenidos.

un mecanismo similar al de las máquinas tragaperras, que proporcionan contenido nuevo o no según una pauta de aleatoriedad. Además, genera una memoria muscular que hace que los propios dedos sean los que actualicen, incluso cuando se ha tomado la decisión de parar.

- Eliminación de las pautas de detención (*stopping cues*)<sup>58</sup>. La existencia de pautas de detención supone la existencia de momentos en los que los usuarios pueden plantearse otras opciones, por lo que los negocios que dependen de la atención tratan de minimizarlas. Al eliminarse las pautas de detención, las actividades nunca terminan. Schüll (2012, caps. 1-3) explica cómo esto se diseña muy cuidadosamente en los casinos: ya se ha mencionado el diseño sin ventanas o ángulos (nota 11), pero se podrían señalar otros. Por ejemplo, en muchas tragaperras, el mecanismo de introducir una moneda y tirar de la palanca ha sido eliminado en favor de pantallas y tarjetas de crédito, que favorecen el juego ininterrumpido; la iluminación oscura y los diseños envolventes también fomentan la pérdida de percepción del entorno, que suele generar pautas de detención. La ausencia de arcos narrativos también supone una ausencia de estas pautas. Si la actividad es repetitiva, sin final o conclusión posible, sin introducción, desenlace, capítulos, intermedios, ni momentos de transformación relevantes, se vuelve una mecánica de final abierto<sup>59</sup> que dificulta el control del agente sobre sus actos. Un ejemplo serían los *infinite scrolling feeds* que mantienen a la gente bajando de manera interminable por los muros de las redes sociales, o con la auto-reproducción de vídeos, que multiplica el fenómeno del *binge watching*<sup>60</sup>. La soledad también se destaca a menudo como característica de las adicciones del comportamiento, principalmente porque supone otra ausencia más de pauta de detención. La existencia de otras personas con otras necesidades y ritmos suele proporcionar momentos de parada y decisión ausentes en soledad.

Para Schüll, la posesión de varios de estos componentes genera un bucle lúdico (*ludic loop*)<sup>61</sup> que funciona muy bien para atraer y mantener clientes en sistemas que se benefician del volumen de uso más que del precio. Un ejemplo sería una aplicación, con mecánicas abiertas, sin pausas de detención, que continúa creando contenidos interesantes y ligeramente impredecibles, a gran velocidad, sin un cierre narrativo, plagado de pequeñas recompensas, para un sujeto en soledad. Los bucles lúdicos poseen una gran capacidad de generar estados de flujo en los que se pierde la percepción del tiempo.

---

58 Un ejemplo de pauta de detención sería la bajada del telón en los teatros o las pausas para la publicidad en la televisión. En las redes sociales, no existen tales pautas, por lo que no existe el momento de parar y decidir, no hay pausa, no hay recuperación de la agencia.

59 Un ejemplo sería servir cervezas que se rellenan infinitamente sin darnos cuenta, de tal modo que uno perdería la noción de cuánto ha bebido y si debe o no parar, sensación que sí tiene cuando debe pedir otra ronda al camarero.

60 *Binge-watching* hace referencia al fenómeno de ver series en tandas de muchos episodios cada vez, se suele traducir por «maratón» o «atracción» de series. El fenómeno ha aumentado drásticamente desde la aparición de las plataformas *online* y de vídeo a demanda. No sólo eso, sino que la información que recopilan estas plataformas genera modelos que se utilizan en producción para maximizar las posibilidades de que un espectador enlace episodios, por medio de la distribución de los segmentos narrativos, los momentos culminantes (*cliff-hangers*), etc. (Matrix, 2014)

61 También habla de ellos Adler (2018, cap 7)

Al contrario que las adicciones a sustancias, las adicciones del comportamiento<sup>62</sup> son muy fáciles de ocultar y pasan desapercibidas durante años. Los comportamientos adictivos no son nada nuevo, pero, a diferencia de tiempos anteriores, la mayoría de las personas disponen ahora de sofisticados dispositivos capaces de proporcionar inmediatamente contenidos diseñados con las características del bucle lúdico las 24 horas del día, incluyendo a los adolescentes, en una etapa especialmente estresante de su vida.

## 5. ¿POR QUÉ FUNCIONAN?

¿Por qué engañan las *fake news*? ¿Por qué pasa la gente más horas de las debidas en redes sociales, incluso si sabe que ello le causa problemas en su vida? ¿Por qué, incluso sabiendo los males de las redes sociales, los individuos caen sistemáticamente en ellas? ¿Cómo es que funcionan tanto el paternalismo libertario como su contrario, analizado en este trabajo? Para explicarlo, hay que comprender un punto fundamental: aunque las grandes empresas tecnológicas defienden sus prácticas utilizando la antropología de la economía clásica, desarrollan sus productos desde los postulados de la economía conductual y la racionalidad limitada; alaban un modelo antropológico que encumbra las capacidades racionales del ser humano, mientras que luego buscan explotar explícitamente sus debilidades mediante todos los recursos de diseño posibles. El principal argumento en contra de intervenir en el mundo de las redes sociales (y en contra de cualquier instancia estructural en general) parte de la asunción de la autonomía, es decir, de que los seres humanos somos, generalmente, racionales y libres, al menos en un grado más que suficiente. La polarización o la adicción (o la pobreza, el paro, el crimen, etc.) son producto de decisiones libres. Es posible que haya algunos males asociados a ciertas prácticas, pero los individuos se bastan para evaluar adecuadamente lo que es mejor para ellos mismos y decidir lo que desean hacer. En la medida en la que no sea producto de coacción (requisito de la libertad), llevar a cabo tal o cual práctica será el resultado del uso de sus facultades racionales (requisito de la racionalidad).

Este modelo se sintetiza habitualmente en el concepto de *homo oeconomicus*, que entiende a los seres humanos como agentes libres y racionales, que se diseñan a sí mismos y, aunque toman sus elecciones en un mundo de restricciones, maximizan la utilidad de sus decisiones usando la información disponible y procesándola apropiadamente dadas sus preferencias y limitaciones. Los errores que cometen los individuos provienen de una carencia de información, racionalidad o autocontrol.

De ser así el ser humano, el mal uso de las POBEA debería ser algo residual y prácticamente desaparecer en el caso de que se proporcionase mejor información al respecto de esas plataformas.

---

<sup>62</sup> Adler (2018) define estas adicciones como comportamientos que la persona no es capaz de resistir y que aunque satisfacen una necesidad a corto plazo, causan daño a largo plazo. Él considera que el factor principal en el desarrollo de una adicción es el contexto en interacción con la capacidad para procesar experiencias negativas como el estrés, la soledad o la angustia; aunque también existen variables biológico-genéticas. Cuando un individuo relaciona un comportamiento (uso de internet, consumo de sustancias) con un resultado atractivo en un contexto de sufrimiento, tenderá a la adicción. Las sustancias o comportamientos que provocan descargas de dopamina en el cerebro no son adictivas en sí mismas hasta que se aprende a usarlas como salvación a los problemas emocionales, como salvavidas para el equilibrio psicológico.



## 5.1 Racionalidad limitada (*bounded rationality*)

Esta antropología comenzó a ser cuestionada desde la economía del comportamiento (*behavioral economics*) y las teorías de la racionalidad limitada, que mostraban cómo los individuos no son puramente racionales, sino que sus decisiones son inconsistentes en el tiempo, se basan a menudo en modelos heurísticos, su comportamiento hacia el riesgo depende de encuadres, puntos de anclaje y disposiciones estructurales, y, a menudo, sus elecciones son contradictorias con un cálculo racional.

Frente al modelo del *homo oeconomicus*, el paternalismo libertario de Thaler y Sunstein parte de una teoría de la racionalidad limitada, que defiende que nunca contamos con información suficiente ni habilidades cognitivas ilimitadas, no nos caracterizamos por el egoísmo y poseemos un autocontrol limitado. El marco, la inercia social, los puntos de partida o cualquiera de las decenas de sesgos cognitivos que tenemos condicionan las decisiones mucho más que un proceso deliberativo racional.

Es una creencia falsa la idea de que la mayor parte de las personas, la mayor parte del tiempo, realizan elecciones de acuerdo con sus mejores intereses, o incluso que los conocen mejor que terceras personas<sup>63</sup>. El ser humano tiende a tomar mejores decisiones en contextos en los que tiene más información y experiencia que en los que no, pero no necesariamente. De hecho, en algunos contextos, los individuos no tienen «preferencias propias» formalmente hablando, ya que estas están fundamentalmente determinadas por la organización del contexto en el que se sitúen. Además, frente a las teorías de la elección racional, las críticas de la racionalidad limitada sostienen que, a menudo, someter a la gente a más elecciones puede hacer que empeoren sus decisiones, porque los individuos no son capaces de manejar grandes cantidades de datos que tienen que ser contrastados entre sí<sup>64</sup>. A menudo, un número de elecciones alto abrumba a los individuos y hace que pospongan la decisión. Por ello, no siempre es mejor para los individuos el ser sometidos a una constante toma de decisiones, incluso aunque sea informada. Ni siquiera el propio proceso de toma de decisiones es siempre una fuente de satisfacción. En ciertas ocasiones lo es, pero en otras, especialmente cuanto más técnica sea la decisión, el tener que tomar partido se vuelve una carga. Un ejemplo habitual es el de los pacientes, que prefieren que los médicos les guíen, antes que tomar decisiones que no comprenden bien sobre su propia salud. Este ser humano, por lo tanto, es altamente sugestionable por el contexto, a menudo carece de preferencias y le cuesta o incluso molesta tomar decisiones. POBEA diseñadas para explotar estas tendencias estarían minimizando la agencia de los individuos y apostando contra ellos.

---

63 Sunstein y Thaler (2013) ponen el ejemplo del jugador de ajedrez novato, que, tomando sus propias decisiones, perderá frente a otro jugador más experimentado precisamente como producto de su inexperiencia y que, si es inteligente, seguirá los consejos de jugadores más expertos para aprender a tomar mejores decisiones. Por ello, la capacidad de la gente para tomar buenas decisiones para sí mismas es, a menudo, una cuestión empírica y no a priori. Añaden en la nota 16 que «Los economistas serán, y deberían ser ignorados si continuamos insistiendo en que es indiscutible que el hecho de comercializar activos constantemente o acumular deudas de consumidor, o volverse un adicto a la heroína debe ser óptimo para la gente que hace esto básicamente debido a que han decidido hacerlo», en definitiva, que el *homo oeconomicus* es un modelo que incurre en una petición de principio.

64 Ejemplo: es más fácil elegir entre 3 opciones que entre 450, sean pantalones, planes de pensiones o apartamentos.

También defienden la racionalidad limitada las teorías del proceso dual, abanderadas principalmente por Kahneman (2012), que presentan una visión del ser humano que procesa a dos niveles, un sistema 1, automático, frecuente, emocional, estereotipado, subconsciente, que genera respuestas intuitivas que funcionan habitualmente en contextos de inmediatez; y un sistema 2, lento, exigente, lógico, consciente y de uso menos habitual, que toma decisiones tras razonar y debe controlar al sistema 1. Este sistema 1 suele dominar en muchas áreas del pensamiento, ya que el sistema 2, por su alto coste de funcionamiento, es mucho más perezoso, y por lo general se complementan. La inteligencia sería una mayor disposición al uso del sistema 2, siendo la causa las decisiones humanas deficientes, a menudo, el uso del sistema 1 para tareas que requieren el 2. Con ello, Kahneman presenta una antropología dual que fragmenta al *homo oeconomicus*<sup>65</sup> y defiende que la racionalidad en sentido clásico no es la característica fundamental del ser humano. Debido a los atajos y sesgos del sistema 1 que nos conducen a errores, lo más conveniente es fomentar y ejercitar el uso del sistema 2 para aumentar la agencia individual y la satisfacción con las propias decisiones, pero las características de las redes sociales ya estudiadas (eliminación de pautas de detención, recompensas aleatorias, retroalimentación rápida, etc.) buscan lo contrario, forzar al uso de los mecanismos del sistema 1.

En ocasiones se ha criticado que estas teorías en el fondo están salvando al *homo oeconomicus*, separándolo del cerebro reptil «malo», pero, en última instancia, preservando el modelo. Parecen querer decir que no somos *homo oeconomicus*, pero nuestros córtex prefrontales (o el sistema 2, etc) sí lo son, por lo que dejan la puerta abierta a que, en el fondo, la responsabilidad sea finalmente individual, por más que las POBEA se programen para explotar todos los sesgos humanos. Se retomará el tema de la responsabilidad individual en la sección 6.

## 5.2 La teoría argumentativa o interaccionista de la razón

Mercier y Sperber (2017) van un paso más allá al criticar tanto la teoría clásica de la razón como las teorías duales. Defienden que el ser humano no es ni un *homo oeconomicus* completo ni lo es a medias, ya que la razón es tan solo un módulo de inferencia intuitiva entre los muchos de que dispone el cerebro humano, cada uno con su estrecho dominio de competencia. Su característica principal es que obtiene conclusiones a partir de razones (motivos) que no son solo intuitivas, sino también reflexivas. Este módulo no se usa todo el tiempo. A menudo, ni siquiera se usa frecuentemente y es un producto evolutivo adaptado especialmente al nicho hipersocial en el que viven los seres humanos. Sus funciones principales son justificar las propias acciones<sup>66</sup> y convencer a otros para que compartan las creencias

---

65 De hecho, no es casual que en la segunda sección de su libro, hable de la comparación entre los humanos según la teoría económica clásica, Econs, racionales y egoístas, y los humanos reales tal cual se comportan en la vida normal.

66 Tarvis y Aaronson (2020, p. 2) «Como seres humanos falibles, todos compartimos el impulso de justificarnos y evitar responsabilizarnos de cualquier acción que resulte dañina, inmoral o estúpida. La mayoría de nosotros nunca estaremos en condiciones de tomar decisiones que afecten la vida y muerte de millones de personas, pero ya sean las consecuencias de nuestros errores triviales o trágicas, a pequeña escala o en un escenario nacional, la mayoría de nosotros encontramos difícil, si no imposible, decir: “Me equivoqué; cometí un terrible error”. Cuanto mayor sean las apuestas (emocionales, financieras, morales), mayor será la dificultad. Va más allá: la mayoría de las personas, cuando se enfrentan directamente a la evidencia de que están equivocadas, no cambian su punto de vista o curso de

propias, su funcionamiento es vago, sesgado y solo muy raramente es objetivo o imparcial. Pero desde el punto de vista de los interaccionistas, los sesgos y vaguedades no son errores, sino características esenciales que sirven para cumplir su función de la manera más eficiente en las interacciones humanas típicas: en lugar de elaborar complejas contraargumentaciones, es evolutivamente más eficaz esperar a que los interlocutores lo hagan. Ahí reside la tercera función de la razón, evaluar las razones de otros, ámbito donde la razón sí suele ser exigente y estricta<sup>67</sup>. Por lo tanto la razón humana no es un instrumento mediante el que se mejora el conocimiento individual del mundo. Además, funciona mejor en grupo (en interacción, dialécticamente), tanto para beneficio de los individuos particulares (primariamente) como del grupo social en general (de forma indirecta), lo que posibilita la cooperación y la moralidad.

Dada esta caracterización de la razón, no es esperable que los individuos, en la soledad de su uso de las POBEA, encuentren las mejores razones para modular su uso de ellas. Más aún, el requisito de la racionalidad deja de ser válido puesto que la razón no es una herramienta modelo para alcanzar la verdad ni la maximización individual, sino para autojustificarse (lo que llevaría a una gran cantidad de personas que hacen un uso patológico de las redes sociales a convencerse de que hacen lo mejor para sí mismas).

La realidad parece que apunta a que el ser humano no es un *homo oeconomicus* que libre pero estúpidamente arruina su vida tomando decisiones sorprendentemente equivocadas, sino que poseemos una racionalidad muy limitada, naturalmente plagada de sesgos (patrones sistemáticos de pensamiento que se desvían del juicio racional<sup>68</sup>) como la aversión a la pérdida, el efecto Dunning-Krueger, el síndrome del impostor, el efecto Barnum, el sesgo de confirmación, el efecto dotación, el sesgo *pro statu quo*, la falacia del jugador, etc<sup>69</sup>. Los diseños tecnológicos sofisticados que aprovechan estos sesgos para fomentar sus modelos de economía de la atención reducen la agencia generando una asimetría tal que merecería una mayor protección de los individuos.

Además, la biología y la psicología han estudiado también los llamados superestímulos (*supernormal stimuli*<sup>70</sup>), estímulos artificiales nocivos que imitan cualidades naturales que normalmente llevarían a

---

acción, sino que lo justifican de manera aún más tenaz. Incluso la evidencia irrefutable rara vez es suficiente para perforar la armadura mental de la autojustificación.» Traducción propia.

67 Por eso somos mucho mejores descubriendo los errores y sesgos de los demás que los propios, algo que no tiene sentido desde el paradigma «intelectualista» de la razón.

68 Para Mercier y Sperber (2017, p. 12) los sesgos son funciones bien ajustadas y no desviaciones de la razón. «[...] los aparentes sesgos que se han descrito como deplorables desperfectos de la razón son en realidad características bien adaptadas a su función argumentativa». Sin embargo, se utilizará la definición clásica porque se está intentando justificar que la visión tradicional de la razón y la autonomía del individuo no son argumento suficiente para avalar la no regulación de la economía de la atención.

69 Una lista exhaustiva disponible en [https://en.wikipedia.org/wiki/List\\_of\\_cognitive\\_biases](https://en.wikipedia.org/wiki/List_of_cognitive_biases) y están maravillosamente agrupados y presentados en: [https://upload.wikimedia.org/wikipedia/commons/6/65/Cognitive\\_bias\\_codex\\_en.svg](https://upload.wikimedia.org/wikipedia/commons/6/65/Cognitive_bias_codex_en.svg)

70 Estudiados originalmente en gaviotas por Nikolaas Tinbergen, que demostró mediante un experimento que las crías de gaviota argéntea prestan más atención a un palo de madera con marcas de alto contraste rojo que a sus propios padres. El rojo es el superestímulo, que en la naturaleza indicaría la garganta de sus padres que porta comida. También encontró que preferían empollar huevos de escayola si eran más grandes o tenían marcas o colores más intensos y que algunas mariposas macho preferían intentar aparearse con réplicas de cartón antes que con hembras

comportamientos adaptativos. Se han estudiado ejemplos con la pornografía en internet (Hilton Jr., 2013), la comida basura e incluso el cine. Es posible que el desarrollo de las POBEA las haya conducido a la creación de superestímulos que aparentan satisfacer nuestras naturales tendencias de entretenimiento, información, conexión social, etc., pero lo hacen proporcionando un sucedáneo perjudicial.

¿Cómo es posible que el diseño de las redes sociales haya llegado a métodos eficientes para la explotación de nuestros sesgos cognitivos, emocionales y biológicos? A través de la principal técnica con la que se refinan los algoritmos: el *A/B Testing*.

### 5.3 A/B Testing

Se denomina *A/B Testing* a una práctica para perfeccionar los algoritmos a través de experimentos aleatorios (aunque pueden ser guiados) sobre dos grupos, uno de control y otro con la variable. Esta es la forma principal mediante la cual las empresas perfeccionan sus algoritmos dirigiéndolos hacia la maximización de la atención y la interacción de los usuarios, favoreciendo los resultados que obtienen más clicks, más tiempo en la aplicación, más comentarios, etc.<sup>71</sup>. Esta es una práctica sistemática de las grandes tecnológicas, en la que testan cientos de variaciones de las características de sus plataformas. Lo que se obtiene a través de estos test son correlaciones eficaces, no causalidades o modelos teóricos que expliquen cómo funcionan, por lo que a menudo es difícil de comprender, incluso para los propios ingenieros y desarrolladores, por qué los algoritmos toman ciertas decisiones<sup>72</sup>. Estos simplemente optimizan lo que sea que funciona a través de la realización masiva de pruebas. En algunas ocasiones, se han llevado a cabo experimentos polémicos, como el que Facebook realizó sobre el contagio emocional con sus usuarios sin comunicárselo<sup>73</sup> (Kramer, 2014; Meyer, 2014). Por supuesto, probar diferentes productos para evaluar cuál es más eficaz no es una técnica innovadora de las POBEA, se lleva practicando desde siempre y lo hacemos todos a pequeña escala. Lo distintivo de las POBEA es la velocidad de testeo, la enorme base sobre la que se experimenta y la capacidad de procesamiento de datos<sup>74</sup>.

---

reales (Barrett, 2010).

71 Un ejemplo sería una web que, para su diseño, prueba dos modelos diferentes en los usuarios que acceden a ella. En uno, los menús, los botones o los colores tienen una configuración y en otro, otra. Aquel modelo que obtenga resultados mejores según las métricas definidas por los programadores se implementa y se repite el proceso.

72 De hecho, a Schüll (2012, cap. 3) le llamaba la atención cómo muchos de los mayores expertos en explotación de las vulnerabilidades humanas no tenían una perspicacia especial sobre la naturaleza humana, más bien al contrario, apenas podían explicar por qué funcionaban los mecanismos que diseñaban. Daban con la fórmula sin entenderla.

73 Facebook mostraba a diferentes usuarios contenido principalmente alegre o triste en sus muros personales y comprobaba que más tarde tendían a producir contenido especialmente positivo o negativo respectivamente.

74 Ejemplo: Un contador de historias de una tribu de cazadores-recolectores seguramente variaba su historia y se quedaba con las variantes que más gustaban y descartaba las que menos, refinando su historia a lo largo del tiempo, una especie de *A/B Testing* ancestral. Su capacidad para encontrar patrones universales que mantuviesen la atención estaba muy limitada y le sería complicado medir qué factor tuvo más influencia en que la historia funcionase de forma diferente de una vez para otra. Facebook o Google pueden llevar a cabo miles de experimentos cada día sobre una base de cientos de millones de personas y disponen de potentes sistemas de IA para analizar todas las variables y así determinar cambios.

#### 5.4 La banalidad del mal digital: La convergencia instrumental y extremismo algorítmico

Los algoritmos así diseñados que determinan el funcionamiento de las POBEA maximizan el tiempo que pasan los usuarios en ellas mediante *A/B Testing* haciendo una suerte de ingeniería inversa del cerebro humano. No hay un complot consciente para perjudicar a la humanidad, sino que son tan solo el resultado de la búsqueda de los diseños más rentables y, desgraciadamente, resulta mucho más rentable la creación de filtros burbuja, la promoción de contenidos que generan miedo o ira, la potenciación de los mecanismos de la adicción ya señalados, etc. Si fuesen más rentable para la economía de la atención los diseños que mejoran la agencia humana, desarrollan virtudes cívicas y promueven el conocimiento, los algoritmos tenderían a ello. Los perjuicios que se derivan de esa búsqueda son debidos a un fenómeno llamado «convergencia instrumental». Esta es la tendencia de agentes inteligentes a perseguir de manera infinita objetivos meramente instrumentales. El ejemplo más icónico lo propuso Bostrom (2003) con su maximizador de clips. Este sería una inteligencia artificial diseñada para fabricar clips de oficina, por lo que si fuera lo suficientemente avanzada, tomaría todas las medidas necesarias para hacer su trabajo. Esto resultaría en que convertiría el planeta Tierra primero y la galaxia después, en la mayor cantidad posible de clips. En el caso de las redes sociales, en la medida en que los algoritmos tienen como objetivo maximizar la atención de los usuarios, su objetivo es encontrar las mejores maneras de atraer a más usuarios durante más tiempo, independientemente de lo que esto implique. Si utilizásemos el mismo principio para el diseño de carreteras, nos encontraríamos con que lo ideal sería llenarlas de violentos accidentes de tráfico, porque son, con diferencia, lo que más atrae la atención de los conductores y pasajeros en la carretera. Afortunadamente, las carreteras tienden a diseñarse con otros objetivos en mente, como la seguridad, la comodidad y la fluidez. Ya que las plataformas digitales poseen las llaves de nuestras actuales ágoras virtuales, quizá no deberían estar orientadas por algoritmos instrumentales e intereses económicos que no responden necesariamente a nuestros intereses individuales o sociales. ¿Sería mejor que los objetivos de las instituciones sociales más importantes (y las POBEA se han convertido ya en una de ellas) sean decididos por la propia sociedad y no por las empresas? Profundizaremos en la cuestión en la sección 7 y en algunas propuestas de solución en la sección 8.

La competición por la atención de los usuarios ha devenido en algunas plataformas en un fenómeno bautizado como extremismo algorítmico, especialmente en Youtube. Esto significa que las recomendaciones que ofrece el portal han ido inclinándose sutil pero inequívocamente hacia tendencias que favorecerían el contenido extremista. Independientemente de dónde se comience, cuando se pasan suficientes horas en Youtube, su algoritmo comenzará a guiar al usuario hacia contenidos cada vez más radicales. La importancia de este problema se pone de manifiesto cuando se observa que el 70% de los mil millones de horas diarias que la humanidad pasa en Youtube proviene de vídeos recomendados por el algoritmo<sup>75</sup>, vídeos que la gente no ha acudido a ver explícitamente, por lo que es alarmante que ese

---

75 Datos sacados del blog oficial de Youtube <https://blog.youtube/news-and-events/you-know-whats-cool-billion->

mecanismo tienda a la radicalización de los usuarios para maximizar el tiempo que pasan en la red. La existencia de tales algoritmos y su forma de operar es una decisión de diseño, no una necesidad esencial de la plataforma, que bien podría no recomendar ningún vídeo en absoluto o estar orientada a promover vídeos que el usuario no llegaría a ver jamás, pero de ser así, los usuarios pasarían muchas menos horas en la plataforma, porque el paso de unos contenidos a otros sería un momento parada y reflexión en el que podrían tomar decisiones como dejar de ver más vídeos, lo que iría en contra de los intereses de la compañía. El extremismo algorítmico ha sido denunciado por investigadores como Munn (2019), periodistas como Naughton, Fisher y Bennhold (2018) y por ex-programadores de la propia plataforma, como Guillaume Chaslot<sup>76</sup>. Todos ellos denuncian que los contenidos extremistas, *fake news* y demás materiales explotan el deseo humano de descubrir conspiraciones y mirar detrás de la cortina. Las teorías conspirativas son excelentes para atraer la atención y mantenernos interesados, para conseguir clicks y horas de visionado, por lo que es un tipo de contenido que el algoritmo tenderá a promocionar y si bien en teoría nosotros podemos elegir cualquier vídeo de los miles de millones que están disponibles, el sistema de recomendación guía a los usuarios a ciertos nichos particulares, por lo que no funciona como un mero recomendador neutral de contenido cualquiera de su red.

Algo más preocupante que los excesos derivados de un algoritmo entrenado hacia fines cuestionables es el hecho de que no solo los ejecutivos de Youtube a menudo conocían estos problemas, sino que sus propios programadores proponían soluciones que eran ignoradas porque reducían el compromiso de los usuarios, lo cual era la prioridad principal de la compañía (Bergen, 2019). En la misma línea, Facebook solo autorizó cambios en el algoritmo a consecuencia de los eventos posteriores a las elecciones presidenciales estadounidenses de 2020, de tal modo que privilegiase fuentes de alta confianza frente a otras poco fiables o políticamente extremistas<sup>77</sup>. El problema de esta intervención es que reducir la prevalencia de estos contenidos dañinos reduce el compromiso de la gente con la red social, por lo que Facebook calibró esa nueva funcionalidad tan solo lo suficiente como para para filtrar el contenido dañino pero dejando pasar parte para que los usuarios vuelvan a por más (Roose et al., 2021).

## **6. LA CUESTIÓN DE LA RESPONSABILIDAD INDIVIDUAL Y EL CASO DE LA ADICCIÓN**

Dado el mecanismo de monetización de la economía de la atención, existe un fuerte incentivo para hacer adictivas sus plataformas. Como ya hemos visto, las redes sociales se benefician económicamente del tiempo que pasan sus usuarios en ellas, por lo que, en ausencia de otros incentivos, lo más racional

[hours/](#) y de <https://www.cnet.com/news/youtube-ces-2018-neal-mohan/>. Probablemente sean superiores en la actualidad, dado que son de 2017 y 2018.

76 La existencia de esta tendencia del algoritmo puede haber sido depurada en la purga que hizo Youtube ante el boycott de anunciantes de 2017 que comenta Bergen en su artículo (2019), porque no pudo ser comprobada en el estudio estadístico de Ledwich y Zaitsev (2019). Ellos se inclinan a pensar que Youtube, en el momento de su estudio, simplemente sugería vídeos dentro del campo habitual del usuario y, si acaso, dirigía el tráfico hacia canales mainstream, si bien reconocen que no tienen la capacidad para testar datos a la escala que sería necesaria y que su categorización política ha sido parcialmente subjetiva. En cualquier caso, discrepan con los autores mencionados y consideran que no se puede decir que el algoritmo de Youtube, en el momento del estudio, radicalice a sus usuarios.

77 Evidenciando que el motor que mantenía la propaganda tóxica, las *fake news* y la extrema polarización hasta ese momento era una decisión de diseño, y que bastaba con cambiar unos pocos aspectos para reducir el daño.

es que traten de maximizar ese tiempo. Estas redes sociales, por defecto, aplican todos sus mecanismos de seducción sobre los usuarios, aunque permiten que estos desactiven muchos de ellos, si así lo desean, a través de los menús de opciones más o menos accesibles. De este modo, los usuarios pueden limitar el número de notificaciones, poner alertas automatizadas cuando pasen un cierto tiempo en la aplicación, etc. Así se traspa al usuario la carga de la responsabilidad, pero, simultáneamente, estas aplicaciones tienen a decenas o miles de ingenieros y psicólogos diseñando cada aspecto de las mismas, perfeccionando los algoritmos, analizando las respuestas del sistema nervioso, los datos que el usuario haya dado ya, sumados a los datos de uso de cientos de millones de usuarios por todo el planeta y sistemas de procesamiento masivo de información. Frente a todo ello, se espera que el usuario sea el que resista la atracción de las redes auto-limitando su uso.

Para entender lo que implica este paradigma de la adicción, se podría pensar en la imagen de Odiseo siendo advertido por Circe de que lo primero que encontrará en su ruta serán las Sirenas:

[...] que a los hombres hechizan venidos allá. Quien incauto se les llega y escucha su voz, nunca más de regreso el país de sus padres verá ni a la esposa querida ni a los tiernos hijuelos que en torno le alegren el alma. Con su aguda canción las Sirenas lo atraen y le dejan para siempre en sus prados; la playa está llena de huesos y de cuerpos marchitos con piel agostada. (Homero, 1993, p. 286)

Lo que le propone Circe para evitarlo es que tape los oídos de sus marineros con cera, que ninguno pueda escuchar la voz de las Sirenas y que él mismo exija que le aten al mástil con cuerdas, de tal modo que pueda disfrutar de su canto y que si, en su locura, exigiera que le suelten o le hagan caso en lo que fuera, le atasen aún con más cuerdas.

Lo que le plantea Circe a Odiseo es la respuesta clásica frente al problema de la adicción: está en las propias manos del individuo esquivar el desastre, la responsabilidad es suya y tiene dos únicas opciones para evitar la ruina. Por una parte, la privación absoluta, llenar nuestros oídos de cera, no probar la droga de la perdición, prohibir las POBEA a los más jóvenes y a nosotros mismos. Por otra, sabiendo que ahora somos racionales, pero que quizá no lo seamos más adelante bajo el efecto seductor de las Sirenas, tomar las medidas para que, incluso al margen de nuestra voluntad, no podamos ya caer en su embrujo o, análogamente, establecer mecanismos para que estas plataformas se cierren automáticamente cuando llevemos tal o cual cantidad de minutos en ellas, no nos envíen notificaciones, etc. La primera propuesta nos priva del placer del canto de las sirenas y la segunda nos rescata cuando hemos perdido la agencia si hemos sido suficientemente precavidos. Las respuestas frente al abuso de las nuevas tecnologías y de las redes sociales que se suele proporcionar son, en su mayoría, de este tipo: se debe mejorar la educación, fomentar la concienciación y la responsabilidad individual. Con ellas se desplaza al sujeto completa e injustamente la responsabilidad del uso: mientras se le conciencia acerca del mismo, se le proporciona acceso a unos dispositivos diseñados para manipularle junto a algunas

herramientas para que prevenga esa manipulación conscientemente diseñada y cuya sofisticación se trata de aumentar cada día. Volviendo a Odiseo, el panorama actual sería como proporcionar a todo el mundo algunas cuerdas a la vez que se les incentiva a pasar por donde están las Sirenas que, cada día, mejoran su canto; y se culpa a los que naufragan en sus costas por su falta de voluntad.

En el caso de las redes sociales y las nuevas tecnologías, a diferencia de las Sirenas de Odiseo, no nos encontramos ante una fuerza sobrenatural, fuera de nuestro control, cuyos perjuicios solo pueden esquivarse evitando por completo su uso o resistirse mediante una autodisciplina feroz, sino ante una creación humana, y, por lo tanto, cabría imaginar a Odiseo proponiendo una tercera posibilidad a Circe: ¿y si transformamos las Sirenas? Podríamos mantener a las Sirenas donde están, pero regularlas, disfrutar de lo atractivo de su canto, pero procurando que estas no se beneficien principalmente con la pérdida de sus oyentes, sino, tal vez, con su desarrollo y bienestar. En términos más contemporáneos, deberíamos estudiar y reformar los mecanismos que generan la adicción y transformar los incentivos que fomentan la persecución de implementar mecanismos de adicción más eficaces. Lo contrario sería propiciar nuestro propio fracaso, como dejar paquetes de cigarrillos gratuitos a las puertas de los colegios o chupitos de tequila de regalo en las gasolineras, confiando en que el buen juicio de los seres humanos será suficiente para evitar la tentación, sorprendiéndonos luego, cuando la salud pulmonar de los jóvenes empeore al ritmo que aumentan los accidentes de tráfico.

### **6.1 El individuo frente a la estructura**

Confinar toda la responsabilidad en el individuo quizá sea pecar de falta de imaginación sociológica. ¿Qué significa esto? La imaginación sociológica es un concepto acuñado por Wright Mills (2000) que describe la capacidad de captar la relación entre la experiencia individual y la sociedad en general, de entender las constricciones que limitan nuestras posibilidades, los nexos que nos unen con la sociedad y propician lo que ocurre en el mundo y en nosotros mismos. Es la comprensión de que la biografía individual es el resultado de un proceso histórico y social que ocurre en un contexto social más amplio que el individuo y que las conductas individuales a menudo son el resultado de ciertas disposiciones sociales. Es la capacidad de conectar los problemas personales a asuntos públicos. La posesión de imaginación sociológica permite hacer análisis más certeros a la vez que ganar autoconciencia sobre las decisiones, para no dejarse arrastrar por las normas sociales u otros factores que modelan nuestras decisiones<sup>78</sup>.

Pogge (2008) diferencia entre dos formas de análisis moral: por una parte el análisis moral interactivo y el análisis moral institucional. El primero defiende que la responsabilidad de una situación injusta tiene que ver con las acciones individuales, mientras que el segundo se fija en cómo está estructurado el

---

<sup>78</sup> Un maravilloso ejemplo lo pone Harari (2016, pp. 72-79) con su descripción del origen de la costumbre de tener un jardín de césped delante de la casa. El conocimiento de su historia nos permite comprender por qué los individuos han llegado a creer que «es bonito» y, si así lo deseamos, liberarnos del «agarrón del pasado» que dirige nuestros futuros.



mundo social en el que se producen los fenómenos para ver cómo dependen estos de las instituciones y la estructura que les rodea. El primer enfoque se utiliza más porque es más fácilmente identificable la trama de relaciones causales que dan lugar a una situación, mientras que la influencia de los elementos estructurales no siempre se deja ver.

Pogge pone el ejemplo de la pobreza que, a menudo, se ha relacionado con la pereza individual, la tendencia a la delincuencia, la drogadicción, el desperdicio de oportunidades, etc; todo lo cual supondría un análisis interactivo. La propia existencia de algún individuo humilde capaz de vencer las circunstancias y encumbrarse a la riqueza era argumento suficiente para considerar que la pobreza era una cuestión de responsabilidad individual. Los análisis institucionales se fijan, en cambio, en las restricciones estructurales a la movilidad social, la limitación de los caminos hacia la riqueza, los mecanismos de los mercados de capital y trabajo, la permanencia de las tasas de desempleo y pobreza, la perpetuación familiar de la riqueza (o pobreza), la influencia en esas tasas desde el rediseño institucional, etc. Los análisis institucionales parten de la idea de que los sujetos toman sus decisiones siempre dentro de marcos y para comprender los fenómenos aparentemente individuales se deben estudiar esos marcos. Las decisiones, aunque parezcan totalmente libres, a menudo reflejan patrones favorecidos por la sociedad. Sin embargo, aquí radica el problema del análisis de muchos fenómenos, ya que para poder hacer análisis morales institucionales se requiere imaginación sociológica, pero muchas veces se carece de ella, se es incapaz de ver lo general en lo particular, el patrón de acción general en las acciones individuales<sup>79</sup>.

Aplicado al caso de las POBEA y los problemas que genera a nivel individual y social, será más interesante hacer un análisis institucional en vez de un mero análisis interactivo centrado en la biografía de los individuos que se vuelven adictos, perjudican su salud, se polarizan hasta la violencia o pierden la noción de la realidad; culpando su falta de autocontrol, su negligente educación o tendencias morales defectuosas. Si tenemos a ingenieros ayudados de supercomputadoras estudiando las debilidades de los cerebros humanos y los caminos de acceso más rápidos a los resortes que necesitan utilizar para la explotación económica de sus aplicaciones, parece que centrarse únicamente en el usuario individual sería, en efecto, una falta de imaginación sociológica<sup>80</sup>.

Desde ese análisis institucional, este trabajo ha pretendido mostrar la influencia de la actual arquitectura de las redes sociales en los fenómenos de polarización, posverdad, adicción, depresión, etc., de una forma análoga a como la arquitectura de los casinos es fundamental para explicar la ludopatía (Schull,

---

79 La diferencia entre estos dos tipos de análisis moral, sostiene Pogge, es la que se ha llegado a identificar con la distinción entre justicia y ética. Aplicado al ámbito de las POBEA, de lo que se tratará es de fomentar la existencia de un sistema justo, sin que ello sea en perjuicio de todas aquellas medidas encaminadas a una ética del autocontrol individual a través de la educación, la concienciación, etc. Es decir, tratar de crear un mar Egeo donde Odiseo y su tripulación puedan transitar con justicia entre Sirenas sin tener que desarrollar una ética heroica para resistirlas.

80 De hecho, muchos de estos problemas ya son conocidos y por ello existen numerosas medidas de regulación para la televisión, especialmente en el horario infantil, que son inexistentes en Youtube.

2012) o la arquitectura de las ciudades es fundamental para explicar muchos de los problemas sociales<sup>81</sup> de la planificación urbanística centrada en el automóvil, y cómo si se diseñan pensando en espacios públicos, comunales, iluminados, etc., desaparecen o se mitigan, aumentando el bienestar de la gente (Jacobs, 2016).

Por lo tanto, si la arquitectura de las redes sociales tiene una cuota de responsabilidad en estos fenómenos, habrá que modificarla. Quedan por contestar dos cuestiones. La primera, ¿se puede delegar en las empresas esa responsabilidad? La segunda, ¿existen propuestas concretas que se pudiesen implementar?

## 7. LA VÍA DE LA RESPONSABILIDAD SOCIAL DE LAS EMPRESAS

Podría plantearse que, en efecto, la actividad de las grandes empresas tecnológicas, además de crear un servicio útil y valioso para muchos, también crea peligrosas externalidades<sup>82</sup>, pero que la solución de esos problemas debería venir desde el ámbito de la responsabilidad de las propias empresas, a través de la herramienta de la Responsabilidad Social de las Empresas (RSE). Esta es la línea del Pacto Mundial (ONU, 2000) y del Libro Verde (CEE, 2001), que se publican en un contexto de globalización, que combina la desregulación de los mercados y las transacciones financieras, la mejora de los transportes y las comunicaciones y un mundo en el que pasamos de economías locales o internacionales mediadas por los estados-nación a economías de escala global. Barañano (2007) argumenta que el propio concepto de

---

81 Hay quienes sostienen, como Jorge Dioni, que el modelo de urbanismo determina incluso la ideología política. (Plaza, 2021)

82 Hablar de «externalidades» ya es cargar políticamente el lenguaje de una determinada manera. Esa visión de la economía, que ha imperado entre muchas posibles, considera que el foco central de la economía son los mercados que operan en base a equilibrios de oferta y demanda que establecen los precios. Así, los precios, la oferta, la demanda y las necesidades de los mercados son las métricas por las que el sistema económico se preocupa y todas las demás variables se denominan «externalidades», especialmente cuando son negativas. Dos ejemplos de externalidades serían la degradación de los barrios cuando una gran cantidad de casas se convierten en pisos turísticos (perdiéndose así el tejido que mantiene los colegios, las tiendas del barrio, la tranquilidad nocturna, las redes vecinales, etc.) o la degradación del medio ambiente que ocurre como resultado de los medios de transporte. Raworth (2017) denuncia que si nos centramos en la actividad económica de esta manera, considerando los pisos de alquiler vacacional o del transporte en términos de oferta, demanda, precio, ganancia individual, etc., y valoramos todos los demás aspectos de esas realidades económicas como «externalidades» secundarias cometemos un profundo error que nos conduce a la conclusión de que la destrucción del planeta o la degradación de los entornos en los que vivimos son tan solo aspectos poco relevantes de la actividad que tratamos. Más aún, critica siquiera que haya que comenzar considerando que el punto de partida en el estudio de una realidad económica sean necesariamente los mercados (los fenómenos de oferta y demanda) y no el entorno en el que se inscriben. Existen ciclos del agua, ciclos del carbono, largas cadenas de biodiversidad y demás aspectos que forman parte fundamental del planeta Tierra que se ven afectados por nuestro uso de los medios de transporte. Igualmente, además de potenciales consumidores u ofertadores de pisos en alquiler, somos residentes, votantes, vecinos, parientes, amigos, padres y madres, miembros de asociaciones y comunidades, etc., por lo que nuestras relaciones con los entornos urbanos en los que vivimos no se reducen a nuestra relación de mercado con los pisos de alquiler. Raworth propone un nuevo modelo para pensar estos problemas llamado *Doughnut economics*, en el que los análisis de las actividades económicas tienen en cuenta a los mercados como solo uno de los factores a considerar, junto con otros aspectos de los objetivos de Desarrollo Sostenible de la ONU como la sanidad, la educación, la equidad social, la seguridad, la ecología, etc. Trasladado esto al mundo de las POBEA, podríamos considerar que Facebook, Twitter, etc., son agentes que generan una oferta para satisfacer una demanda y cuyo éxito se basa en monetizar el tráfico en base a los anuncios que ponen. Sin embargo siguiendo esta perspectiva, podríamos dejar de considerar como meras externalidades los problemas individuales y sociales que estas redes sociales generan, en su configuración actual, y pasar a considerarlos como parte fundamental de su existencia, de tal modo que establezcamos regulaciones que desincentiven sus comportamientos actuales.

RSE no es unívoco, y la disputa por su significado<sup>83</sup> y su carácter multiescalar, multidimensional y multidireccional son una expresión de las características de nuestro tiempo, donde el modelo de estado-nación se ha visto confrontado y a menudo superado por la globalización y el despliegue de las corporaciones a lo largo de cadenas de producción intercontinentales. Los gigantes tecnológicos exceden la jurisdicción de cualquier estado, por lo que conseguir que trabajen bajo un paradigma de desarrollo sostenible sería más eficaz que una maraña de legislaciones diversas en los distintos estados. Además, sostienen los defensores, la inversión en estos fines sociales más allá del lucro se puede plantear en términos de oportunidad, ventaja competitiva o inversión estratégica.

Sin embargo, no parece una vía fructífera. A la vista los casos ya mencionados (Mozur, 2018; Bergen, 2019; Fisher, 2019; Mac, 2021; Roose et al., 2021), parece claro que los gigantes tecnológicos se muestran renuentes a tomar las medidas necesarias cuando estas afectan directamente a sus fuentes de ingresos, y solo se mueven significativamente cuando los anunciantes amenazan con retirarse o las instituciones públicas legislan. En otras ocasiones alegan que existe una buena voluntad, pero simplemente son incapaces de poner freno a estos problemas. Resulta llamativo que la misma inteligencia sea simultáneamente capaz de levantar esas empresas pero no encuentre solución a los desperfectos que generan. En cualquier caso, de ser así, dada la magnitud de los problemas analizados, la incapacidad para atajarlos no es un eximente, sino un motivo mayor de preocupación.

La RSE, a menudo, se ha mostrado como una mera herramienta de marketing empresarial, valiosa para las compañías en la medida en que mejora su reputación, pero subordinada a otra responsabilidad aún mayor, la que tiene la empresa para con los accionistas. Sin embargo, el marketing no debería ser la herramienta con la que se gestionan los problemas que generan las empresas. Maira Vidal (2013) va más allá y habla de la RSE como un instrumento típico del neoliberalismo para detener o, incluso, hacer retroceder las regulaciones ya existentes, privatizando, en la práctica, el derecho; delegando todas las cuestiones de la responsabilidad debida de las empresas a códigos que se caracterizan por ser voluntarios, unilaterales y autorregulados; sustituyendo códigos jurídicos y normativos o contratos legales con castigos y contrapartidas por códigos morales; relegando la elaboración de todos los instrumentos (códigos, memorias, etc.) al ámbito privado de cada empresa, otorgando con ello capacidad de las empresas para elegir qué derechos son importantes o en qué condiciones los van a respetar; y cooperando solo con empresas de verificación o auditoría privadas elegidas por la propia empresa a ser auditada, en vez de por un servicio público de inspección.

Sandel (2012) también critica las consecuencias éticas y políticas de esta deriva del rol de los mercados en la que se ha pasado de tener economías de mercado a convertirnos en sociedades de mercado, es decir, hemos pasado de usar los mercados como herramientas valiosas y eficaces para organizar

---

83 Barañano destaca tres concepciones principales: La RSE como plus normativo, la RSE como cumplimiento de la ley y filantropía y la RSE desde la perspectiva de las «partes interesadas». El modelo más habitual es el primero, aunque desde el mundo empresarial también se defiende el segundo, comparten las características de ser voluntarios, unilaterales y autorregulados.

aspectos de la actividad productiva a un modo de vida en el que los valores del mercado son centrales en todos los aspectos de la vida humana, relegando las posibilidades humanas de decidir juntos qué tipo de vida valoramos y queremos a simplemente nuestra capacidad adquisitiva en un mercado que ofrece todo de manera (superficialmente) neutral. Sandel considera que debemos cuestionar el rol y el alcance de los mercados, el tipo de cosas que deben poder comprarse y venderse y el tipo de cosas que deben estar sometidas a otros valores no de mercado. Debemos retomar la pregunta sobre qué mundo queremos. ¿Deseamos que el mundo sea únicamente el resultado de las decisiones de compañías que buscan maximizar sus beneficios económicos y, en el caso que nos atañe, maximizando la manipulación de nuestros sistemas nerviosos con una creciente asimetría de poder sobre la que no recae ninguna responsabilidad o preferimos que haya regulaciones que limiten las posibilidades de las POBEA y orientarlas a aumentar la agencia individual y las virtudes cívicas, como el fortalecimiento de las comunidades, los derechos humanos, la pluralidad y la democracia? La responsabilidad de los dirigentes de las empresas es, primariamente, con los accionistas de las mismas y solo de forma secundaria deben preocuparse de ser benefactores de la humanidad, en la medida en que eso puede producir mejores resultados económicos. No es maldad, sino la esencia del sistema económico en el que operan estas empresas. Dadas las reglas de juego vigentes, para las empresas es mucho más rentable aprovechar nuestros sesgos y los reflejos de nuestra amígdala y del sistema límbico que generan nuestras respuestas más agresivas, más impulsivas, que hacen que sigamos volviendo una y otra vez a la red social. Si es cierto, como decía Edward Wilson en la cita que abre este trabajo, que los humanos tenemos un cerebro paleolítico, instituciones medievales y tecnología de dioses, y que esta diferencia se va a ir exacerbando, tendremos que hacer lo posible por ajustar las instituciones al auge tecnológico; ya que paralizarlas cediendo la responsabilidad del control a quienes desarrollan tecnologías potencialmente destructivas es suicida. Por lo tanto, parece prudente proponer algunas medidas con las que tratar de controlar, desde fuera del ámbito de la RSE, los problemas que generan estas compañías, animando, por supuesto, a que estas vayan más allá con sus propios códigos de buenas prácticas si lo desean.

## **8. PROPUESTAS**

La situación actual no es inevitable, sino tan solo el producto de incentivos mal orientados y una escasa o inexistente regulación sobre la economía de la atención. Para mitigar los problemas que genera, se podrían proponer diferentes soluciones. Aquí se presentan algunas que van en orden creciente de intervencionismo:

### **8.1 Propuestas de aumento de la información**

Las aplicaciones, portales, redes sociales, etc., que contengan ciertos mecanismos como los estudiados anteriormente (recompensas aleatorias, *fast feedback*, notificaciones, ludificación, etc.), podrían incluir advertencias sobre los efectos negativos que producen, de manera similar como ocurre con el tabaco.

También se podría fomentar el aumento de la transparencia, por ejemplo, forzando a las plataformas a indicar el número de veces que un contenido ha sido recomendado, o qué proporción de las visitas de un contenido proviene de recomendaciones y cuál de búsquedas directas. De este modo se podrían diagnosticar y mitigar más rápidamente los problemas a los que conduce la convergencia instrumental (por ejemplo, la ya mencionada tendencia de Youtube a fomentar la pedofilia o el extremismo), y así promover la responsabilidad de las plataformas sobre sus recomendaciones. Un ejemplo sería la plataforma sin ánimo de lucro llamada *Algotransparency*<sup>84</sup>, que trabaja monitorizando y publicando el impacto de los algoritmos de Google (sugerencias de autocompletado), Youtube (vídeos recomendados), Facebook (paneles con más interacción) y Twitter (*trending topics*), por ejemplo, mostrando cuáles son los contenidos más recomendados por las diferentes redes y en qué cantidades.

Estas propuestas no son suficientes, pero serían un paso en la buena dirección. Que los usuarios conozcan los riesgos ayuda a tomar decisiones mejor informadas y una sociedad que sepa hacia dónde dirigen la atención los algoritmos puede decidir si le gusta cómo lo hacen y si respetan los valores y necesidades de esta. ¿Fomentan contenidos que promueven la anorexia o las conspiraciones? ¿En qué proporción? Estas medidas proporcionarían herramientas de diagnóstico con las que generar un debate adecuado, por lo que es una buena pieza del puzzle, pero no la única necesaria. Ya se ha comentado la insuficiencia del modelo epistemológico y antropológico del *homo oeconomicus*, que supone decisiones racionalmente ideales ante la información disponible, por lo que el aumento de la información no es suficiente. Los adictos no son necesariamente individuos poco informados, en ocasiones, todo lo contrario<sup>85</sup>. La idea de que información completa conduce a una agencia total proviene de ese modelo antropológico que considera que todo ser humano debería revisar racionalmente sus dinámicas de comportamiento negativo a la luz de la nueva información, y descubrirlas evidentemente contrarias a su interés, por lo que el adicto o el radical informado aparecerían como humanos disfuncionales. Sin embargo, estamos más bien ante seres humanos normales cuyos resortes cognitivo-emocionales naturalmente sesgados y limitados se han explotado, y las adicciones y demás se nos muestran, más bien, como síntomas de una estructura disfuncional.

## 8.2 Propuestas desde el paternalismo libertario

Para ellas seguimos la tipología ya presentada en la sección anterior sobre paternalismo libertario.

- Paternalismo mínimo: las plataformas podrían traer, por defecto, desactivadas las opciones que generan la mayor parte de los problemas: *infinite scroll*, sistema de recomendación de vídeos, reproducción automática, notificaciones, visualización de *likes* y seguidores, etc.; ofreciendo la posibilidad de activarlas. También podrían incluirse sistemas automáticos de información del número de horas pasadas en la aplicación activados por defecto.

---

84 <https://www.algotransparency.org/>

85 Pensemos en los fumadores, perfectamente conscientes de los efectos del tabaco. Schüll (2012) destaca que muchos de los ludópatas a los que entrevistaba eran matemáticos o ingenieros/técnicos de máquinas tragaperras.

- Elecciones activas requeridas: las plataformas podrían incluir un sistema de menús obligatorio al darse de alta que obligase al usuario a decidir sobre las opciones anteriormente descritas o, incluso, sobre el tiempo diario/semanal que desea pasar en ellas, el tipo de uso que quiere que la aplicación le fomente, etc.
- Las restricciones de procedimiento: las plataformas podrían dificultar la activación de las opciones anteriores (situándolas en un sistema de menús incómodo, reseteando la decisión cada día y obligando a hacerlo de nuevo, etc.).
- Restricciones sustantivas: además de dificultar la activación, las plataformas podrían permitirlo solo cierta cantidad de veces en periodos de tiempo estipulados.
- Rechazo de la libertad de elección: implicaría la prohibición completa de las redes sociales a grupos de personas, como menores de edad, igual que se hace con el alcohol u otras drogas.

### 8.3 Cambios al incentivo de los algoritmos

Una opción diferente, aunque con características comunes con el paternalismo libertario, sería fomentar que los algoritmos se optimicen para encontrar otras soluciones que no sean la maximización del tiempo ante la pantalla. Se los podría entrenar para detectar los comportamientos negativos y transformarlos, en vez de explotarlos<sup>86</sup>. Por ejemplo, se podría tratar de dificultar el uso de las aplicaciones cuando se detecten patrones de uso compulsivo, creando distancias, introduciendo fricciones. Un estudio (Holte y Ferraro, 2020) muestra que tan solo cambiando a blanco y negro sus dispositivos se reducía significativamente el tiempo que pasaban en ellos<sup>87</sup>. Se podría también ralentizar la velocidad de conexión o la fluidez de las aplicaciones de manera incrementada pero aleatoria, para que los usuarios no puedan predecir cuándo ocurrirá, a partir de una cantidad de minutos de uso recomendada, recuperando así pautas de parada para que el usuario pueda plantearse el uso que hace. Las redes sociales también podrían detectar la pérdida de control del usuario, detectar su soledad y susceptibilidad de caer en dinámicas depresivas y mostrar activamente qué amigos están conectados o cerca y sugerir iniciar conversación<sup>88</sup>. Todas estas opciones podrían ser desactivables. Los algoritmos podrían entrenarse en función del *feedback* de los usuarios, por ejemplo, mostrando estadísticas de uso mensuales y preguntádoles si están satisfechos con la cantidad y calidad del tiempo empleado, si se arrepienten o no<sup>89</sup>, si les está ayudando a aprender cosas nuevas y mantenerse conectados, fomentando

86 En la nota 17 ya vimos todo lo que son capaces de predecir ya hoy en día por el comportamiento *online* del usuario.

87 Y con ello, los resultados negativos como ansiedad, depresión, trastornos del sueño y baja satisfacción vital.

88 Para proteger la privacidad, se podría implementar a través de una autorización explícita en la que cada usuario indicase un número de usuarios a los que, en caso de riesgo de comportamientos depresivos (propios o ajenos), se les indique la necesidad de contactar.

89 A menudo, muchos usuarios se sorprenden de la cantidad de horas que han pasado enlazando vídeos de Youtube que, retrospectivamente, no les han aportado nada. Esto no ocurre necesariamente con todo el tiempo pasado ante pantallas, como el que se pasa viendo alguna buena película, jugando con amigos, etc. Algunos proponen la creación de test de arrepentimiento (Eyal, 2020), un test que suspenderían las *apps* cuyos usuarios lamenten haber utilizado (en general, en cantidad de horas, de usos concretos, etc.). Estos test de arrepentimiento serían especialmente buenos

así el aumento de la agencia de los usuarios.

#### 8.4 Cambios legales significativos

Legislativamente, podrían fomentarse cambios de diferente calado. Una posibilidad sería la restricción de la economía de la atención, limitando la cantidad de tiempo que las aplicaciones pueden monetizar la atención de los usuarios, por ejemplo, permitiendo mostrar tan solo anuncios durante los primeros 30 minutos de uso cada día, eliminando así uno de los incentivos para diseñarlas de la forma más adictiva posible. Una modificación más agresiva sería la obligación de pagar reparaciones, realizando estudios sociológicos que relacionasen las problemáticas sociales con el uso de este tipo de aplicaciones y calculando tasas de compensación, análogas a las tasas de emisiones por perjudicar el medio ambiente que se imponen (o deberían) a los mayores emisores globales. Se podría por último, prohibir completamente la economía de la atención, obligando a la eliminación de todos los metadatos y castigando severamente a todas las aplicaciones que los recopilen en el futuro, elaboren modelos predictivos y segmenten contenidos. Se eliminaría así la opción de que los individuos pudieran dar permiso para que se recopilen y vendan sus datos. Esta medida podría apoyarse no solo en argumentos paternalistas sobre los beneficios para la autonomía y la libertad que supondría para los individuos, incluso contra su voluntad, sino en la posibilidad de que, como defiende Véliz (2020, cap. 3), la privacidad sea un bien colectivo, aunque estemos tratando de datos personales. Al igual que en las cuestiones ecológicas, las acciones de unos individuos repercuten sobre otros. Nuestra interdependencia mutua en cuestiones de privacidad implicaría que no existe un derecho moral individual por el cual las personas pueden querer vender libre y legítimamente sus propios datos, porque los datos de cada uno contienen ya en sí datos de otros<sup>90</sup>. Para Véliz, la privacidad es colectiva en dos aspectos: por un lado, una brecha en la privacidad de un individuo puede facilitar violaciones de la privacidad de otros individuos; por otro, las consecuencias de la pérdida de privacidad se experimentan colectivamente.

Una propuesta diferente sería cambiar el estatus legal de los gigantes tecnológicos, introduciendo un deber de cuidado o deber fiduciario:

En inglés existe un término, *fiduciary*, cuya traducción al castellano fiduciario no tiene el mismo

---

para derribar algunos de los mitos del *homo oeconomicus*, ya que una de las bases morales del libre mercado es que satisface la autonomía creando riqueza y permitiendo la distribución de recursos en función de los deseos de la gente. Si existiera una abundante evidencia empírica de que los sujetos están crecientemente dedicándose a tareas de las que se arrepienten, el paradigma de la racionalidad limitada (y las protecciones que implica) ganaría peso.

<sup>90</sup> Véliz ilustra esta cuestión con el ejemplo de los *kits* de ADN para el trazado de los árboles genealógicos. Un individuo puede desear comerciar con su información genética, pero el porcentaje de genes específicamente individuales que hay en cada persona es del 0,1%. Esas similitudes con el código de los demás (especialmente con aquellos más cercanos genealógicamente a nosotros) permite que, en la medida en la que se comparten características, lo que se aplica con éxito a unos individuos se puede reproducir con éxito en otros. Además, cruzando información de ADN con otros datos, incluso estando anonimizada, permite identificar con gran precisión a grandes cantidades de personas, incluso aunque no hayan proporcionado su propio ADN a la base de datos. Otro ejemplo más sencillo podría ser que un individuo, al ceder los datos de su teléfono a alguna compañía, está exponiendo también muchos datos de sus contactos.

sentido<sup>91</sup>. Un *fiduciary* es una persona que mantiene un vínculo legal y ético basado en una relación de confianza con una o más partes y está sujeto a deberes fiduciarios (*fiduciary duties*), que restringen los comportamientos posibles de manera mucho más estricta; son los niveles más altos de cuidado exigibles. Se espera un alto grado de lealtad del fiduciario a la persona con la que contrae el deber. Esta relación se contrae cuando en la dinámica entre las partes hay una gran asimetría de poder<sup>92</sup>. Algunos ejemplos de este tipo de relación serían la que mantiene un sacerdote con aquellos que se confiesan, un abogado con sus clientes o un médico con sus pacientes. El *fiduciary duty* regula estas relaciones para proteger a los más débiles de esta relación de poder asimétrica. Un médico, un abogado o un sacerdote pueden querer aprovechar su posición de poder con la información obtenida en el curso natural de su trabajo para chantajear o simplemente para recomendar soluciones que no sean las mejores para la salud, situación legal o espiritual de las personas a las que atienden, sino las que les reportan más beneficios. Esta asimetría es esencial para el desempeño de su trabajo. Ni el abogado, ni el médico ni el sacerdote pueden hacer bien su labor si no recopilan abundante información sensible de sus clientes/pacientes/feligreses, pero, al hacerlo, la relación resultante ya no puede ser la de un contrato entre iguales. Por ello, el *fiduciary duty* establece límites muy estrechos a lo que pueden hacer las partes con poder con esa información, que siempre ha de ser en beneficio de la otra parte.

En el escenario actual, la relación contractual que existe entre los usuarios y las empresas tecnológicas es de igualdad. El usuario accede a los servicios y con ello ciertos derechos sobre sus datos a la empresa para que los utilice como considere (en conformidad con la ley). El uso principal que hacen estas empresas con esa información es crear perfiles sobre los usuarios (o modificar su comportamiento) para que otros agentes usen la atención, pagando por el espacio, el tiempo y la segmentación personalizada de los usuarios. Los usuarios están utilizando POBEA diseñadas por miles de ingenieros, psicólogos y demás expertos, perfeccionadas con decenas de millares de experimentos para encontrar las mejores formas de aumentar su tiempo en la red, a la vez que se acumulan inmensas cantidades de información personal que aumentan las capacidades predictivas del comportamiento de los usuarios, lo que genera una enorme asimetría de poder entre usuarios y compañías, por ello podría considerarse que estas plataformas deberían estar sometidas a un deber de cuidado o fiduciario. Este obligaría a las POBEA a actuar solo en beneficio de sus usuarios y se podría detallar en qué consiste ese beneficio de forma similar a como se hace con abogados, médicos o sacerdotes, con las métricas y restricciones que resulten relevantes<sup>93</sup>. Si los médicos, los abogados o los sacerdotes están sometidos a deberes fiduciarios, ¿no cabría utilizar esta asimetría, exponencialmente mayor, como base para regulación que

91 En castellano, un fiduciario es una persona encargada de un fideicomiso y de la propiedad de bienes que lo integran.

92 Por eso es típica de las herencias. El fiduciario debe ser leal y respetar al máximo los intereses de la parte que lo contrató, ya que la asimetría es máxima por la muerte de esta y su incapacidad de defenderse.

93 Evaluaciones positivas en los test de arrepentimiento (mencionados en la nota 89), disminución de síntomas de enfermedades mentales, ahorro económico, etc. Estas métricas podrían estar moduladas por decisiones de los usuarios de acuerdo con POBEA diseñadas según los criterios apuntados en la sección 8.2, de tal modo que se castigue severamente a aplicaciones que fomentan la adicción o generen depresión y arrepentimiento de uso, pero en menor medida si el usuario ha recorrido cada menú de la aplicación para activar todos y cada uno de los resortes adictivos (notificaciones, *infinite scrolls*, etc.), en los que constan advertencias e información suficiente.



dificulte al máximo la explotación de las debilidades de los usuarios? Sometiendo a las empresas tecnológicas a un deber de cuidado, se podría regular de tal modo que estas solo pudieran utilizar la información que recolectan sobre sus usuarios para el bien de estos, y no para venderles productos, ideas, generar adicción, etc. Mientras aprovechar esta asimetría no sea ilegal o no conlleve grandes repercusiones fiscales, las compañías tendrán un enorme incentivo económico para hacerlo y, todavía más, una obligación hacia sus accionistas para ello.

## 9. CONCLUSIONES

El desarrollo evolutivo que configuró los cerebros de los seres humanos es un proceso lento que nos ha dado capacidades asombrosas, pero también una racionalidad limitada, con numerosos sesgos cognitivos y emocionales. La tecnología que hemos sido capaces de desarrollar avanza a un ritmo incomparablemente superior a la adaptación evolutiva y nos ha vuelto crecientemente capaces de manipular los resortes de la mente humana para diferentes fines. Gran parte de las instituciones creadas desde la Modernidad están basadas en concepciones liberales sobre la importancia del individuo, su autonomía y su responsabilidad: el cliente tiene razón, el votante sabe lo que le conviene, hay que confiar en los sentimientos y en el corazón de cada uno, etc. Sin embargo, permitir que la tecnología actual haga ingeniería inversa de nuestras capacidades y explote las debilidades para beneficiarse de un modelo económico concreto pone en riesgo múltiples aspectos de nuestras vidas. Crear modelos predictivos cada vez mejores que puedan modificar la conducta con mayor efectividad mientras nos sitúan en burbujas informativas aisladas anula la autonomía de los individuos, las bases de la democracia y somete todo el sistema social a unos intereses económicos que no tienen como objetivo nuestro mayor interés. Los puntos desarrollados a lo largo del trabajo están abiertos a discusión, y futuras líneas de investigación deberían profundizar en el análisis de la relación entre el uso de POBEA y la polarización política o los trastornos psicológicos individuales, el estudio de instancias en las que la arquitectura de un entorno determine (o module decisivamente) el comportamiento de los individuos, así como en el desarrollo de modelos antropológicos más ajustados a los datos que proporcionan las investigaciones académicas y la realidad social; pero si la fundamentación teórica ha sido lo suficientemente sólida, cabe sostener que desde el punto de vista de las tres corrientes éticas principales (éticas deontológicas, consecuencialistas y de la virtud), parece que nos enfrentamos a una crisis política, moral y filosófica: la progresiva reducción de la agencia individual y el aumento de la predicción del comportamiento humano acaba con la autonomía (deontologismo); nos dirige a sociedades más polarizadas, violentas, asustadas, deprimidas y ansiosas, es decir, menos felices (consecuencialismo) y, por último, promueve formas de vida que nos alejan de la clase de personas que querríamos ser, de modelos de excelencia, inclinándonos hacia vicios, adicciones y nos aleja de vidas prudentes y felices (éticas de la virtud). Esta situación no es necesaria o inevitable, sino que es producto de decisiones de diseño que se pueden transformar, apelando a la responsabilidad de quienes diseñan esos mecanismos, pero, sobre todo, cambiando el sistema de incentivos que hace que el panorama actual

sea el más racional dadas las condiciones existentes, lo que se puede hacer a través de leyes e intervenciones decididas colectivamente que busquen fomentar los valores que consideramos socialmente importantes. Este trabajo ha defendido que la situación de parasitismo libertario y la asimetría de poder existente entre las POBEA y los usuarios exige medidas correctivas, que deberían ir en la línea de la limitación de la economía de la atención para ser verdaderamente eficaces. Además, deberían ser apoyadas siempre por otras disposiciones que apuesten por el aumento de la transparencia y la desactivación por defecto de los mecanismos de monitorización y fomento de la adicción de estas plataformas, junto con el apoyo a aquellas que utilicen algoritmos con incentivos socialmente positivos (no arrepentimiento, conexión, mejora de la salud, etc.). Estas medidas no son una lista exhaustiva, no son las únicas posibles, cabrían muchas otras de diferentes grados de ambición regulativa, pero son coherentes con el marco teórico desarrollado y pueden servir como punto de partida para un debate sobre la profundidad de la reforma necesaria para mitigar el problema. En este sentido, el objetivo de alcanzar un cierto consenso político es realista, porque aquellos que valoran la protección y la regulación de los excesos del mercado que proporcionan los Estados podrían estar muy cerca de los que anteponen el valor de la autonomía del individuo y la libertad, ya que la arquitectura actual de las POBEA arremete contra el bienestar de los individuos a la vez que minimiza su agencia. Volviendo a la cita inicial, si es cierto que poseemos una tecnología casi divina, deberemos actualizar nuestras instituciones para controlarla y que no sobrepase los límites de nuestras debilidades paleolíticas, poniendo en riesgo todas nuestras fortalezas.

La verdad última y oculta del mundo es que es algo que construimos nosotros, y que fácilmente podríamos hacerlo de otra manera.

— David Graeber, *La utopía de las normas: De la tecnología, la estupidez y los secretos placeres de la burocracia*

## Referencias bibliográficas

- Angwin, J., Larson, J., Mattu, S., y Kirchner, L. (2016, 13 mayo). Machine Bias. *ProPublica*. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Almagro Holgado, M.; Osorio Mancilla, J.; Villanueva Fernández, N. (2021). *Las Torres de Lucca*.10(19), pp. 43-57
- Alter, A. (2018). *Irresistible*. Paidós.
- Arendt, H. (1998). *Los orígenes del totalitarismo* (G. Solana, Trad.). Taurus.
- Barañano Cid, M. (2009). Contexto, concepto y dilemas de la responsabilidad social de las empresas transnacionales europeas: una aproximación sociológica/Context, concepts and dilemmas in corporate social responsibility for European multinational companies: a sociological approach. *Cuadernos de Relaciones Laborales*, 27(1), 19.
- Barber, B. M., Lee, Y. T., Liu, Y. J., y Odean, T. (2014). *The cross-section of speculator skill: Evidence from day trading*. *Journal of Financial Markets*, 18, 1-24.
- Barrett, D. (2010). *Supernormal stimuli: How primal urges overran their evolutionary purpose*. WW Norton & Company.
- Bergen, M. (2019, 2 abril). YouTube Executives Ignored Warnings, Letting Toxic Videos Run Rampant. *Bloomberg*. <https://www.bloomberg.com/news/features/2019-04-02/youtube-executives-ignored-warnings-letting-toxic-videos-run-rampant>
- Bostrom, N. (2003). *Ethical issues in advanced artificial intelligence. Science fiction and philosophy: from time travel to superintelligence*, 277-284.
- Cardenal, A. S., Aguilar-Paredes, C., Cristancho, C., y Majó-Vázquez, S. (2019). Echo-chambers in online news consumption: Evidence from survey and navigation data in Spain. *European Journal of Communication*, 34(4), 360-376.
- Carr, N. (2008, julio). Is Google Making Us Stupid? *The Atlantic*. <https://www.theatlantic.com/magazine/archive/2008/07/is-google-making-us-stupid/306868/>
- Chandler, D. (1995). *The act of writing: a media theory approach*. University of Wales.
- Choy, S. K. (2015). Retail clientele and option returns. *Journal of Banking & Finance*, 51, 26-42.
- Comisión de las Comunidades Europeas (2001). Libro verde: fomentar un marco europeo para la responsabilidad social de las empresas. *CEE*.
- Denham, H. (2021, 15 enero). These are the platforms that have banned Trump and his allies. *Washington Post*. <https://www.washingtonpost.com/technology/2021/01/11/trump-banned-social-media/>
- Diuk, C. (2014, 14 febrero). The Formation of Love. *Facebook Data Science*. <https://www.facebook.com/notes/10158928005273415/>
- Dunbar, R. I. (1998). *Grooming, gossip, and the evolution of language*. Harvard University.
- Dunbar, R. I. (2016). Do online social media cut through the constraints that limit the size of offline social networks?. *Royal Society Open Science*, 3(1).
- Eisenstat, Y. (2016, 17 junio). Former Intel. Official: American Hate Is a Bigger Threat Than Foreign Terrorism. *Time*. <https://time.com/4370375/american-hate-national-security/>
- Eurobarometer, S. (2017). 88. Autumn 2017. *Report. European Citizenship*.
- Eyal, N. (2014). *Hooked: How to build habit-forming products*. Penguin.
- Eyal, N. (2020, 23 marzo). Want to Design User Behavior? Pass the 'Regret Test' First. *Nir and Far*. <https://www.nirandfar.com/regret-test/>
- Fisher, M., y Bennhold, K. (2018, 7 septiembre). As Germans Seek News, YouTube Delivers Far-Right Tirades. *The New York Times*. <https://www.nytimes.com/2018/09/07/world/europe/youtube-far-right-extremism.html>
- Fisher, M., y Taub, A. (2019, 4 junio). On YouTube's Digital Playground, an Open Gate for Pedophiles. *The New York Times*. <https://www.nytimes.com/2019/06/03/world/americas/youtube-pedophiles.html>
- Fung, B. (2021, 11 enero). Parler app has now been booted by Amazon, Apple and Google. *CNN*. <https://edition.cnn.com/2021/01/09/tech/parler-suspended-apple-app-store/index.html>
- García-Pelayo, G. G. (2004). *La Fabulosa Historia De Los Pelayos / The Fabulous Story of the Pelayos*. Penguin Random House.
- Gascón, D. (2018, 18 junio). 10 apuntes sobre la posverdad. *Letras Libres*. <https://www.letraslibres.com/espana-mexico/politica/10-apuntes-sobre-la-posverdad>
- Grind, K., Schechner, S., McMillan, R., y West, J. (2019, 15 noviembre). How Google Interferes With Its Search Algorithms and Changes Your Results. *WSJ*. <https://www.wsj.com/articles/how-google-interferes-with-its-search-algorithms-and-changes-your-results-11573823753>
- Haidt, J., y Allen, N. (2020). Scrutinizing the effects of digital technology on mental health. *Nature* 578, 226-227 (2020). <https://doi.org/10.1038/d41586-020-00296-x>

- Hamilton, J. M. (2019, 16 abril). How Yellow Journalism Sparked the Spanish-American War. *History*. <https://www.nationalgeographic.com/history/history-magazine/article/yellow-journalism-role-spanish-american-war>
- Harari, Y. N. (2016). *Homo Deus: Breve historia del mañana*. Debate.
- Harari, Y. N. (2018). *21 lecciones para el siglo XXI*. Debate.
- Harris, T. (2021, 12 enero). Cómo romper el negocio basado en nuestra atención de las «Big Tech». *MIT Technology Review*. <https://www.technologyreview.es/s/13075/como-romper-el-negocio-basado-en-nuestra-atencion-de-las-big-tech>
- Heid, M. (2019, 14 marzo). Depression and Suicide Rates Are Rising Sharply in Young Americans, New Report Says. This May Be One Reason Why. *Time*. <https://time.com/5550803/depression-suicide-rates-youth/>
- Hern, A. (2020, 1 julio). Facebook ad feature claims to predict user's future behaviour. *The Guardian*. <https://www.theguardian.com/technology/2018/apr/16/facebook-ad-feature-predict-future-behaviour>
- Hernández, J. A. (2000, 12 mayo). El clan de los Pelayo gana la apuesta. *EL PAÍS*. [https://elpais.com/diario/2000/05/12/ultima/958082401\\_850215.html](https://elpais.com/diario/2000/05/12/ultima/958082401_850215.html)
- Hilton Jr, D. L. (2013). Pornography addiction—a supranormal stimulus considered in the context of neuroplasticity. *Socioaffective Neuroscience & Psychology*, 3(1), 20767.
- Holte, A. J., y Ferraro, F. R. (2020). True colors: Grayscale setting reduces screen time in college students. *The Social Science Journal*, 1-17.
- Homero (1993). *Odisea* (trad. Pabón, J. M.). Gredos.
- Jacobs, J. (2016). *The death and life of great American cities*. Vintage.
- Jacobson, L., Sherman, A., Valverde, M., McCarthy, B., y Greenberg, J. (2021, 8 enero). PolitiFact - A day of crisis at the US Capitol, fact-checked. *PolitiFact*. <https://www.politifact.com/article/2021/jan/06/day-crisis-us-capitol-fact-checked/>
- Juste, M. (2021, 10 febrero). La pandemia dispara el uso de las redes sociales, un 27% más que hace un año. *Expansion*. <https://www.expansion.com/economia-digital/innovacion/2021/02/10/6022c89de5fdea59448b459b.html>
- Kahneman, D. (2012). *Pensar rápido, pensar despacio*. Debate.
- Kambhampati, S., Megerian, C., y Kannan, S. (2021, 7 enero). How a pro-Trump mob stormed the D.C. Capitol. *Los Angeles Times*. <https://www.latimes.com/politics/story/2021-01-06/how-pro-trump-mob-stormed-us-capitol>
- Klebnikov, S. (2020, 17 junio). 20-Year-Old Robinhood Customer Dies By Suicide After Seeing A \$730,000 Negative Balance. *Forbes*. <https://www.forbes.com/sites/sergeiklebnikov/2020/06/17/20-year-old-robinhood-customer-dies-by-suicide-after-seeing-a-730000-negative-balance/>
- Kramer, A. D., Guillory, J. E., y Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24), 8788-8790.
- Krasnova, H., Wenninger, H., Widjaja, T., y Buxmann, P. (2013). Envy on Facebook: a hidden threat to users' life satisfaction? *Bern Open Repository and Information System* <https://boris.unibe.ch/47080/>
- Kross, E., Verduyn, P., Demiralp, E., Park, J., Lee, D. S., Lin, N., ... y Ybarra, O. (2013). Facebook use predicts declines in subjective well-being in young adults. *PloS one*, 8(8).
- LaFrance, A. (2020, 15 diciembre). Facebook Is a Doomsday Machine. *The Atlantic*. <https://www.theatlantic.com/technology/archive/2020/12/facebook-doomsday-machine/617384/>
- Ledford, H. (2019). Millions of black people affected by racial bias in health-care algorithms. *Nature*, 574(7780), 608-610.
- Ledwich, M., y Zaitsev, A. (2019). Algorithmic extremism: Examining YouTube's rabbit hole of radicalization. *arXiv preprint arXiv:1912.11211*.
- Lenore, V. (2020, 7 septiembre). La factura de la guerra cultural de Iván Redondo. *Vozpópuli*. [https://www.vozpopuli.com/altavoz/cultura/guerra-cultural-ivan-redondo\\_0\\_1389461205.html](https://www.vozpopuli.com/altavoz/cultura/guerra-cultural-ivan-redondo_0_1389461205.html)
- Lógica, S., y UCM (2017). "I Estudio sobre el Impacto de las Fake News en España". Recuperado el 07 de Mayo de 2018, de <https://d3vjcw65af87t.cloudfront.net/novacdn/EstudioPescanova.pdf>
- Mac, R. (2021, 20 enero). Facebook Is Showing Military Gear Ads Next To Insurrection Posts. *BuzzFeed News*. <https://www.buzzfeednews.com/article/ryanmac/facebook-profits-military-gear-ads-capitol-riot>
- Maira Vidal, M. del M. (2013). La responsabilidad social empresarial como parte del proyecto político y económico neoliberal. *Lan Harremanak - Revista De Relaciones Laborales*, (28). <https://doi.org/10.1387/lan-harremanak.10537>
- Malik, K. (2017, 26 febrero). Not post-truth as too many 'truths'. *Pandaemonium*. <https://kenanmalik.com/2017/02/05/not-post-truth-as-too-many-truths>
- Martín, I. (2021, 8 enero). MAGA Patriots, cómo construir un movimiento al margen de la realidad. *ctxt.es | Contexto y Acción*. <https://ctxt.es/es/20210101/Politica/34669/maga-patriots-alt-right-donald-trump-twitter-isabel-martin.htm>

- Martínez, G. (2018, 1 julio). Fake and fury. *ctxt.es* | *Contexto y Acción*.  
<https://ctxt.es/es/20180627/Firmas/20480/posverdad-fake-news-carta-noticias-Guillem-Martinez.htm>
- Matrix, S. (2014). The Netflix effect: Teens, binge watching, and on-demand digital media trends. *Jeunesse: Young People, Texts, Cultures*, 6(1), 119-138.
- Matsa, K. E., y Shearer, E. (2018). News use across social media platforms 2018. Pew Research Center, 10.
- McIntyre, L. (2018). *Post-truth*. MIT Press.
- McLuhan, M. (1994). *Understanding media: The extensions of man*. MIT.
- Mercier, H., y Sperber, D. (2017). *The enigma of reason*. Harvard University.
- Meyer, R. (2014, 9 agosto). Everything We Know About Facebooks Secret Mood Manipulation Experiment. *The Atlantic*. <https://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/>
- Mozur, P. (2018, 15 octubre). A Genocide Incited on Facebook, With Posts From Myanmar's Military. *The New York Times*. <https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html>
- Muiruri, M. F. G. (2021). "How Cyber Activity Influences Depression and Anxiety in Teenagers". *Walden Dissertations and Doctoral Studies*. 10521. <https://scholarworks.waldenu.edu/dissertations/10521>
- Müller, K., y Schwarz, C. (2020). Fanning the flames of hate: Social media and hate crime. *Journal of the European Economic Association*. <https://doi.org/10.1093/jeea/jvaa045>
- Munn, L. (2019). Alt-right pipeline: Individual journeys to extremism *online*. *First Monday*, 24(6). <https://doi.org/10.5210/fm.v24i6.10108>
- Naughton, J. (2018, 23 septiembre). However extreme your views, you're never hardcore enough for YouTube. *The Guardian*. <https://www.theguardian.com/commentisfree/2018/sep/23/how-youtube-takes-you-to-extremes-when-it-comes-to-major-news-events>
- Newton, C. (2019, 25 febrero). The secret lives of Facebook moderators in America. *The Verge*. <https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona>
- North, A. (2014, 5 noviembre). When Novels Were Bad For You. *Op-Talk*. <https://op-talk.blogs.nytimes.com/2014/09/14/when-novels-were-bad-for-you/>
- Orben, A., y Przybylski, A. K. (2019). The association between adolescent well-being and digital technology use. *Nature Human Behaviour*, 3(2), 173-182.
- Organización de las Naciones Unidas (2010). Pacto Mundial – Responsabilidad Social empresarial. <https://www.pactomundial.org/> y [https://www.pactomundial.org/wp-content/uploads/2018/02/Flyer-New-Strategy-GC-2018\\_20180126.pdf](https://www.pactomundial.org/wp-content/uploads/2018/02/Flyer-New-Strategy-GC-2018_20180126.pdf)
- Paniagua, A. (2021, 25 julio). El suicidio se convierte en la primera causa de muerte entre los jóvenes. *Heraldo - heraldo.es*. <https://www.heraldo.es/noticias/nacional/2021/07/25/el-suicidio-se-convierte-en-la-primer-causa-de-muerte-entre-los-jovenes-1508739.html>
- Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin.
- Pennock, R. T. (2010). The postmodern sin of intelligent design creationism. *Science & Education*, 19(6-8), 757-778.
- Phillips, A. (2021, 6 enero). Mitch McConnell's forceful rejection of Trump's election 'conspiracy theories'. *Washington Post*. <https://www.washingtonpost.com/politics/2021/01/06/mitch-mcconnells-forceful-rejection-trumps-election-conspiracy-theories/>
- Plaza, A. (2021, 23 mayo). Jorge Dioni: «El modelo PAU de ciudad dispersa podría resumirse en: sálvese quien pueda». *ElDiario.es*. [https://www.eldiario.es/economia/jorge-dioni-modelo-pau-ciudad-dispersa-resumirse-salvese-pueda\\_1\\_7956971.html](https://www.eldiario.es/economia/jorge-dioni-modelo-pau-ciudad-dispersa-resumirse-salvese-pueda_1_7956971.html)
- Pogge, T. (2008). ¿Qué es la justicia global?. *Revista de economía institucional*, 10(19), 99-114.
- Popper, N. (2020, 8 julio). Robinhood Has Lured Young Traders, Sometimes With Devastating Results. *The New York Times*. <https://www.nytimes.com/2020/07/08/technology/robinhood-risky-trading.html>
- Rangappa, A., y Mercieca, J. (2020, 11 julio). The Ancient Greeks' Guide to Rejecting Propaganda and Disinformation. *Zócalo Public Square*. <https://www.zocalopublicsquare.org/2020/06/07/fake-news-disinformation-propoganda-truth-rhetoric-twitter-president-trump-ancient-greek-philosophers-socrates-plato-truth/ideas/essay/>
- Raworth, K. (2017). *Doughnut economics: seven ways to think like a 21st-century economist*. Chelsea Green Publishing.
- Roland, Z. (2016, 28 abril). A Socialist's Campaign for California Governor, And the Dirty Hollywood Politics That Sunk It. *KCET*. <https://www.kcet.org/shows/lost-la/the-socialist-who-won-a-democratic-primary-and-the-dirty-hollywood-politics-that-sunk-his-campaign>
- Roose, K., Isaac, M., y Frenkel, S. (2021, 7 enero). Facebook Struggles to Balance Civility and Growth. *The New York Times*. <https://www.nytimes.com/2020/11/24/technology/facebook-election-misinformation.html>
- Rosenbaum, E. (2019, 4 abril). IBM artificial intelligence can predict with 95% accuracy which workers are about to quit their jobs. *CNBC*. <https://www.cnbc.com/2019/04/03/ibm-ai-can-predict-with-95-percent->



[accuracy-which-employees-will-quit.html](#)

- Roy, A., Nikolitch, K., McGinn, R., Jinah, S., Klement, W., y Kaminsky, Z. A. (2020). A machine learning approach predicts future risk to suicidal ideation from social media data. *NPJ digital medicine*, 3(1), 1-12.
- Sadeh-Sharvit, S., Fitzsimmons-Craft, E. E., Taylor, C. B., y Yom-Tov, E. (2020). Predicting eating disorders from Internet activity. *International Journal of Eating Disorders*, 53(9), 1526-1533.
- Sandel, M. J. (2012). *What money can't buy: the moral limits of markets*. Penguin
- Schüll, N. D. (2012). *Addiction by design*. Princeton University.
- Segenni, S. (2019) *False notizie... fake news e storia romana: Falsificazioni antiche, falsificazioni moderne*. Mondadori Education. P. 74 y ss: Francisco Pina Polo, Noticias falsas, desinformación y opinión pública en la Roma republicana.
- Shaer, M. (2014, abril). What Emotion Goes Viral the Fastest? *Smithsonian Magazine*.  
<https://www.smithsonianmag.com/science-nature/what-emotion-goes-viral-fastest-180950182/>
- Shriver, S. K., Nair, H. S., y Hofstetter, R. (2013). Social ties and user-generated content: Evidence from an online social network. *Management Science*, 59(6), 1425-1443.
- Simon, H. A. (1996). Designing organizations for an information-rich world. *International Library of Critical Writings in Economics*, 70, 187-202.
- Snyder, T. (2017). Sobre la tiranía: veinte lecciones que aprender del siglo XX. Galaxia Gutenberg.
- Stanley, J. (2020). *How fascism works: The politics of us and them*. Random House Trade Paperbacks.
- Stevens, H., Santamariña, D., Rabinowitz, K., Uhrmacher, K., y Muyskens, J. (2021, 7 enero). How Republicans in Congress voted on counting Joe Biden's electoral college win. *Washington Post*.  
<https://www.washingtonpost.com/graphics/2021/politics/congress-electoral-college-count-tracker/>
- Sunstein, C. R., y Thaler, R. H. (2003). Libertarian paternalism is not an oxymoron. *The University of Chicago Law Review*, 1159-1202.
- Taub, A., y Fisher, M. (2018, 22 agosto). Facebook Fueled Anti-Refugee Attacks in Germany, New Research Suggests. *The New York Times*. <https://www.nytimes.com/2018/08/21/world/europe/facebook-refugee-attacks-germany.html>
- Thaler, R. H., y Sunstein, C. R. (2003). Libertarian paternalism. *American economic review*, 93(2), 175-179.
- Thompson, N. (2019, 23 abril). Tristan Harris: Tech Is 'Downgrading Humans.' It's Time to Fight Back. *Wired*.  
<https://www.wired.com/story/tristan-harris-tech-is-downgrading-humans-time-to-fight-back/>
- Twenge, J. M. (2017). *iGen: Why today's super-connected kids are growing up less rebellious, more tolerant, less happy--and completely unprepared for adulthood--and what that means for the rest of us*. Simon and Schuster.
- Twenge, J. M., Haidt, J., Joiner, T. E., y Campbell, W. K. (2020). Underestimating digital media harm. *Nature Human Behaviour*, 4(4), 346-348.
- UN News / Health. (2019, 10 diciembre). Measles 'misinformation campaigns' through social media, fuel rising. UN News. <https://news.un.org/en/story/2019/12/1052801>
- Valenzuela, S., Halpern, D., y Katz, J. E. (2014). Social network sites, marriage well-being and divorce: Survey and state-level evidence from the United States. *Computers in Human Behavior*, 36, 94-101.
- Véliz, C. (2020). *Privacy is power*. Random House.
- Wang, Y., y Kosinski, M. (2018). Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of personality and social psychology*, 114(2), 246.
- Whittaker, Z. (2020, 12 mayo). Facebook to pay \$52 million to content moderators suffering from PTSD. *TechCrunch*. <https://techcrunch.com/2020/05/12/facebook-moderators-ptsd-settlement/>
- Wright Mills, C. (2000). *The sociological imagination*. Oxford University.
- Young, D. G. y Miller, J. M. (2021, August 27). Political Communication en *Oxford Handbook of Political Psychology (Third Edition)*. <https://doi.org/10.31219/osf.io/mwdu>
- Zuboff, S. (2020). *La Era del Capitalismo de la Vigilancia*. Paidós.