

Economía conductual y desinformación sanitaria: más allá de los *nudges*



Por José María Abellán Perpiñán,
Fernando Ignacio Sánchez Martínez
y Manuel García Goñi

El análisis de una disciplina cada vez más consolidada que fusiona principios y evidencias provenientes de la economía y la psicología. El impacto de la desinformación en la decisión de las personas, los mecanismos para mitigar el daño y el diseño de estrategias e intervenciones (*nudges*) que ayudan a tomar mejores decisiones en la atención sanitaria. El rol de la economía del comportamiento a la hora de fortalecer las competencias de la ciudadanía

1. Introducción

El mercado sanitario es uno de los objetos de estudio más atractivos para los economistas. Desde luego lo resulta para los firmantes de este artículo, que compartimos la condición de economistas de la salud, pero creemos no equivocarnos si generalizamos su

atractivo al común de los analistas económicos poseedores de una mínima curiosidad intelectual.

Aunque son variadas las razones que justifican esta afirmación, entre ellas, por ejemplo, la importancia que reviste para las sociedades el análisis de la regulación del aseguramiento y provisión sanitaria

* **José María Abellán Perpiñán** es licenciado en Ciencias Económicas y Empresariales y doctor en Economía. En la actualidad ocupa el puesto de catedrático de Economía Aplicada en el departamento del mismo nombre en la Universidad de Murcia, España.

Fernando Ignacio Sánchez Martínez es doctor en Ciencias Económicas y Empresariales y catedrático de Economía Aplicada en la Universidad de Murcia, España.

Manuel García Goñi es Ph.D. in Economics por Boston University y catedrático de Economía Aplicada en el departamento de Economía Aplicada, Estructura e Historia de la Universidad Complutense de Madrid, España.

—uno de los pilares del estado de bienestar—, en este artículo abordamos una motivación muy diferente. Para ello, comencemos recordando que el mercado, o mejor, los mercados sanitarios se caracterizan por adolecer de múltiples fallos a cuyo estudio lleva dedicándose la ciencia económica desde antes incluso de recibir ese calificativo, cuando aún era economía política, allá por el siglo XIX.

Como bien sabemos, los supuestos de competencia perfecta que pueblan los libros de texto introductorios en economía no dejan de ser abstracciones ideales y, por tanto, irreales del comportamiento de los agentes que operan en los mercados. En el caso específico de los “mercados” sanitarios pueden identificarse muchas desviaciones de estos supuestos ideales.

Así, es notorio que en este ámbito abundan los precios regulados, muy poco transparentes por lo general; también hay externalidades, tanto positivas como negativas; prestaciones que poseen en algún grado características de bienes públicos, como las vacunas; y, además, como está sólidamente diagnosticado, la información que poseen los agentes que operan en estos mercados no solo es imperfecta, sino que habitualmente se halla distribuida asimétricamente entre ellos, lo cual puede ocasionar notables distorsiones, como la existencia de variaciones injustificadas en la práctica clínica.

Si instituciones, empresas, clínicos y pacientes fuesen agentes perfectamente racionales, todos estos fallos del mercado simplemente no existirían. Sin embargo, como ha sido ampliamente establecido tanto en el ámbito experimental como en el mundo real, la racionalidad de los seres humanos es limitada, no perfecta (Simon, 1955). Esta es la premisa fundamental que subyace a los análisis que realiza la economía conductual o del comportamiento.

Esta disciplina, hoy día completamente consolidada, fusiona principios y evidencias provenientes de la economía y la psicología, con el objeto de diseñar

estrategias e intervenciones conocidas como *nudges* (Thaler y Sunstein, 2008; 2021), y traducidas al español como «empujones» o «acicates» (véase, por ejemplo, Abellán y Jiménez Gómez, 2020), que nos ayudan a tomar mejores decisiones, también sobre nuestra salud y la de otros.

A esto podría objetarse que desde hace mucho tiempo se adoptan medidas como, por ejemplo, los impuestos sobre el tabaco o, incluso, la prohibición de fumar en espacios públicos, que persiguen la misma finalidad que los “empujones” saludables: prevenir la enfermedad, fomentar la salud. Sin embargo,

los *nudges* se diferencian de estas otras intervenciones “duras” en que no obligan, no son coercitivos. Por eso se dice que forman parte de una suerte de paternalismo libertario o “blando” (Sunstein, 2014), ya que un *nudge*, para serlo realmente, nunca debe imponerse.

La fundamentación psicológica de los *nudges* radica en el mecanismo dual de procesamiento de la información que poseemos los humanos (Kahneman, 2011). Fallamos porque muchas de nuestras decisiones son impulsivas o automáticas, propias de un sistema de pensamiento rápido y, por tanto, inconsciente. Por esa razón —aunque no es la única, el entorno social también importa, y mucho— el fumador recalci- trante, aun cuando se prometa a sí mismo que no volverá a fumar, cada día vuelve a incurrir en el mismo error. El reverso reflexivo de este cortocircuito

mental sería el sistema de pensamiento lento, analítico y, por tanto, consciente. Este sistema deliberado de toma de decisiones es en el que nos basamos para decidir que es una buena inversión suscribir una póliza de seguro sanitario (lo cual, luego, paradójicamente, como ya señalara Arrow, 1963, puede tener un “efecto adverso” indeseado, en forma de una demanda sanitaria excesiva).

En resumen, los *nudges* deben respetar en cualquier caso la libertad de elección del individuo y propiciar que sus decisiones estén alineadas con

“La economía del comportamiento no ofrece una vacuna única contra la desinformación, sino que aporta distintas intervenciones dirigidas a prevenir su aceptación, limitar su difusión o reducir sus consecuencias. Su eficacia aumenta cuando se combinan medidas sobre el individuo, el mensaje, la plataforma y la institución, y cuando se parte de un diagnóstico de las barreras concretas”.

sus intereses de largo plazo, aun cuando eso pueda entrañar la colisión con su placer o satisfacción en el corto plazo. Obsérvese que la legitimidad de este planteamiento reside precisamente en que el “cortoplacismo” no deja de ser, según la economía del comportamiento, manifestación de sesgos cognitivos como, por ejemplo, el sesgo del presente, por el que tendemos a otorgar un peso desproporcionado a los beneficios presentes que nos reporta el hábito malsano -el placer de fumar, beber o engullir ultra-procesados mientras vemos series de televisión sin solución de continuidad- frente a los beneficios futuros que se desprenderían de un estilo de vida más saludable -mejor salud cardiovascular y respiratoria, menor peso y un modo de vida más activo-.

Reparamos en este artículo en un problema relativamente poco analizado en el dominio de la economía conductual, como es el origen, mecanismos causales e impacto de la desinformación -más coloquialmente denominada “bulos” o *fake news* en inglés- relativa a cuestiones concernientes a la salud y la asistencia sanitaria.

La expansión acelerada del uso de las tecnologías de la información durante las últimas décadas ha hecho que circulen diariamente, de manera inmediata, una gran cantidad de informaciones de toda índole sobre salud y sanidad; informaciones, las más de las veces, sin verificar, que son compartidas masivamente en redes sociales, pudiendo originar en ocasiones auténticos problemas de salud pública, como tuvimos ocasión de comprobar durante la pandemia de la COVID-19.

En este artículo profundizamos, por tanto, en cómo la desinformación relacionada con la salud impacta en las decisiones de las personas, no solo porque la información compartida no sea correcta, sino fundamentalmente porque se canaliza de manera que apela directamente a nuestro sistema inconsciente, lo que desafía nuestra capacidad de respuesta ante los bulos, sean deliberados o no. Por este motivo, para combatir la desinformación no basta con corregir o regular los contenidos. Es necesario intervenir sobre la exposición, la atención, la credibilidad, las capacidades cognitivas y la propia arquitectura digital.

En las páginas siguientes, partimos de una explicación conductual de cómo se difunde la desin-

formación y se propicia la adhesión a esta para, a continuación, presentar un catálogo de posibles intervenciones conductuales -de los *nudges* a los *hyper-nudges*- conducentes a la prevención y atenuación del impacto nocivo de los bulos sanitarios. Sería ilusorio pensar que estas herramientas conductuales pueden eliminar por completo la desinformación, pero sí que pueden mitigarla, preservando la autonomía, el pluralismo y la confianza de la ciudadanía.

2. Del bulo al daño

La desinformación sanitaria no es un problema meramente comunicativo. Es un determinante de decisiones que pueden afectar a la salud individual, a la salud pública y al funcionamiento de los sistemas sanitarios. Aunque en el lenguaje corriente ambos fenómenos se agrupen bajo la palabra “bulo”, conviene distinguir entre información errónea (*misinformation*), que es falsa o engañosa con independencia de la intención de quien la difunde, y desinformación (*disinformation*), creada o compartida deliberadamente para engañar o causar daño. La intención importa a la hora de atribuir responsabilidades y diseñar la respuesta, pero no se requiere su concurrencia para que se produzcan consecuencias perjudiciales (OMS, 2024).

La cadena causal que conduce desde el mensaje hasta el daño comprende, al menos, seis eslabones: exposición, atención, aceptación, recuerdo, difusión y conducta. Ninguno es automático: una persona puede recordar una afirmación sin creerla o compartirla sin considerarla cierta. Sin embargo, la repetición, la plausibilidad narrativa, la emoción y las señales sociales facilitan el avance. Además, el desmentido no siempre elimina la influencia del contenido inicial, un fenómeno conocido como “efecto de influencia continuada” (Ecker et al., 2022).

Los daños pueden clasificarse en cuatro niveles. El primero es clínico: rechazo o retraso de intervenciones eficaces, adopción de remedios inútiles o peligrosos, automedicación, incumplimiento terapéutico y demora en buscar asistencia. Por ejemplo, durante la pandemia se difundió en varios países latinoamericanos el dióxido de cloro como supuesto tratamiento frente a la COVID-19, pese a carecer de eficacia de-

mostrada, lo que motivó una advertencia expresa de la OPS (Organización Panamericana de la Salud, 2020; Mostajo-Radji, 2021). En un experimento realizado en Reino Unido y Estados Unidos, la exposición a desinformación sobre las vacunas contra la COVID-19 redujo entre 6,2 y 6,4 puntos porcentuales la proporción de personas inicialmente dispuestas a vacunarse con certeza; los mensajes con apariencia científica resultaron especialmente influyentes (Loomba et al., 2021).

El segundo nivel es poblacional. Cuando una conducta genera externalidades -vacunarse, usar antibióticos de forma responsable o seguir medidas de protección ante un brote-, el perjuicio trasciende a quien toma la decisión. Una disminución pequeña de la adherencia puede traducirse, a escala, en más contagios, resistencias antimicrobianas, complicaciones y muertes evitables. El tercero es organizativo y económico: consultas inducidas por falsas alarmas, pruebas innecesarias, presión sobre líneas de atención, tiempo clínico dedicado a corregir bulos y desvío de recursos desde actividades de mayor valor. El

cuarto es institucional: deterioro de la confianza en profesionales, autoridades, ciencia y medios. Este daño acumulativo puede ser más persistente que el provocado por un bulo concreto, porque reduce la eficacia de comunicaciones futuras y debilita la capacidad de respuesta ante nuevas crisis.

La relación no es unidireccional. Una comunicación institucional tardía, contradictoria o excesivamente categórica puede crear el vacío que ocupa la desinformación; también las experiencias negativas con el sistema sanitario alimentan la desconfianza. La respuesta no puede limitarse a proporcionar más datos a una ciudadanía pretendidamente irracional. El problema combina falsedades, incertidumbre real, incentivos de las plataformas y contextos sociales que hacen verosímiles determinadas narrativas.

3. Por qué creemos y compartimos

La economía del comportamiento ayuda a explicar por qué personas razonables aceptan o difunden afirmaciones de escasa calidad. En entornos saturados



Fuente = Imagen generada con Chat GPT

de información, la atención es limitada y decidir qué creer exige tiempo, conocimientos y esfuerzo. Las personas recurren, por tanto, a atajos mentales –también llamados heurísticas– que suelen ser útiles, pero que también pueden explotarse deliberadamente.

Uno de esos atajos es la fluidez de procesamiento: tendemos a considerar más familiar, plausible y cierta la información que se comprende con facilidad. La repetición refuerza esa sensación y genera el llamado “efecto de verdad ilusoria” (Hasher et al., 1977). No hace falta recordar dónde se leyó una afirmación; basta con que, al reaparecer, resulte conocida. De ahí la desventaja estructural de la corrección: el bulo suele ser breve, rotundo y repetible, mientras que la evidencia científica acostumbra a incorporar matices, probabilidades y condiciones. Un análisis de más de 92.000 contenidos halló que la desinformación era, en promedio, más fácil de leer, menos diversa léxicamente y mucho más proclive al lenguaje emocional negativo que las noticias fiables (Carrasco-Farré, 2022).

La disponibilidad constituye otro mecanismo. Sobreestimamos la frecuencia o probabilidad de suce-

sos que vienen rápidamente a la memoria (Tversky y Kahneman, 1973). El relato vívido de un supuesto efecto adverso de un fármaco puede pesar más que la evidencia sobre los beneficios de millones de dosis administradas. A ello se añade el sesgo de negatividad (Rozin y Royzman, 2001): las amenazas captan la atención con especial rapidez, algo adaptativo en muchos contextos, pero rentable para contenidos alarmistas. El encuadre también importa (Tversky y Kahneman, 1981). No produce la misma impresión decir que una intervención “reduce el riesgo un 25%” que indicar que lo reduce del 0,16% al 0,12 %, aunque ambas formulaciones describan el mismo resultado. La presentación relativa puede magnificar la percepción del beneficio o del daño y, a la inversa, las variaciones absolutas pueden minimizarla. Este sesgo afecta no solo a pacientes, sino también a periodistas, profesionales y decisores.

El sesgo de confirmación favorece la búsqueda e interpretación de información congruente con las creencias previas (Nickerson, 1998). Este mecanismo puede verse reforzado por el razonamiento mo-



Fuente = Imagen generada con Chat GPT

tivado: cuando una conclusión resulta identitaria, emocional o ideológicamente deseable, las personas evalúan las pruebas de manera selectiva para poder alcanzarla o preservarla (Kunda, 1990). Sin embargo, no toda aceptación de información falsa puede atribuirse a una defensa ideológica consciente. Una parte relevante responde a que, en el momento de decidir, la persona no se detiene a evaluar la exactitud. La arquitectura de las redes activa otras metas: entretener, expresar indignación, ayudar a los allegados, ser el primero en informar o reafirmar la pertenencia al grupo. Desplazar momentáneamente la atención hacia la veracidad mejora el discernimiento de lo que se comparte. Esto sugiere que muchas personas no pretenden difundir falsedades, sino que, al decidir si comparten una publicación, atienden principalmente a otros atributos —como su novedad, impacto emocional o afinidad con sus creencias— y no a su exactitud. (Pennycook et al., 2021; Pennycook y Rand, 2022).

Las heurísticas de autoridad y consenso también son decisivas: tendemos a considerar más creíble una afirmación cuando procede de una fuente percibida como experta o cuando aparece respaldada por muchas personas, sin evaluar necesariamente la solidez de las pruebas que la sustentan (Chaiken, 1980; Eagly y Chaiken, 1993; Sundar, 2008). El problema surge cuando las credenciales pueden fingirse, una bata blanca sustituye a la competencia acreditada o el número de “me gusta” se interpreta como prueba. Las plataformas agregan señales de popularidad sin distinguir si proceden de personas, cuentas automatizadas o campañas coordinadas. Así, la “viralidad” se convierte en una señal endógena de credibilidad: un contenido parece importante porque circula, y circula porque parece importante (efecto de arrastre o *bandwagon*).

Las emociones no son un ruido que se añade al razonamiento; orientan la atención, la memoria y la acción. Miedo, ira, sorpresa e indignación incrementan la disposición a compartir, sobre todo cuando el mensaje ofrece un culpable y una explicación sencilla. La desinformación convierte incertidumbre en certeza narrativa: identifica una causa, atribuye una intención y promete control. Las teorías conspirati-

vas satisfacen necesidades epistémicas —entender-, existenciales —reducir la sensación de impotencia- y sociales —preservar una identidad positiva del grupo-. Por eso, responder únicamente con cifras puede fracasar si no se atienden la emoción y la preocupación que sostienen la creencia.

Crear y compartir, además, son conductas distintas. Se puede reenviar un mensaje “por si acaso”, para advertir, provocar o señalar afiliación, aunque se dude de su contenido. Cada reenvío aumenta la exposición y la familiaridad de terceros. La evidencia sobre Twitter mostró que las noticias falsas se difundían más lejos, más deprisa y con mayor profundidad que las verdaderas, debido principalmente al comportamiento humano y no solo a los robots (Vosoughi et al., 2018). La novedad y la reacción emocional ayudan a explicar esa ventaja.

Finalmente, las creencias se forman en redes sociales concretas. La confianza suele depositarse en familiares, líderes comunitarios, profesionales próximos o referentes religiosos y políticos. Cuando una corrección amenaza la identidad del grupo o procede de una institución percibida como ajena, puede ser rechazada sin que el problema sea falta de comprensión. En salud, la vulnerabilidad informativa se distribuye de forma desigual: menor alfabetización sanitaria y digital, barreras lingüísticas, precariedad, experiencias de discriminación y acceso deficiente a fuentes fiables pueden aumentar la exposición y reducir la capacidad de contrastar. La desinformación es, por tanto, también un problema de equidad.

4. Inteligencia artificial, autoridad sintética e *hypernudges*

La inteligencia artificial (IA) generativa modifica la economía de la desinformación. Reduce drásticamente el coste de producir textos, imágenes, audios y vídeos plausibles; permite generar innumerables variantes de un mismo relato; traduce y adapta el mensaje a registros locales; y automatiza su difusión y respuesta. La novedad no reside únicamente en fabricar un contenido falso de mayor calidad, sino en combinar escala, velocidad, personalización y conversación.

Los modelos generativos pueden aportar información sanitaria comprensible y facilitar el acceso, pero

también producir errores con tono seguro, referencias inexistentes o argumentos persuasivos a favor de afirmaciones falsas. En experimentos controlados, participantes no expertos tuvieron dificultades para distinguir tuits escritos por personas de los generados por GPT-3, y este pudo producir desinformación más convincente que la redactada por humanos (Spitale et al., 2023). La OMS ha advertido de que el uso sanitario de grandes modelos multimodales exige gobernanza, validación, transparencia, protección de datos y supervisión humana (OMS, 2025a).

Denominamos aquí “autoridad sintética” a la apariencia de competencia, imparcialidad y conocimiento que emerge de una respuesta fluida, personalizada y formalmente bien construida, aunque no exista detrás una fuente responsable que responda de su veracidad. En una conversación, el sistema puede recordar lo dicho, aparentar empatía, anticipar objeciones y ajustar el nivel de lenguaje. Estas cualidades son valiosas, pero también pueden elevar indebidamente la confianza. La persuasión deja de ser un mensaje estático dirigido a un público medio y se convierte en una interacción capaz de aprender qué argumento funciona con cada usuario.

La personalización automatizada enlaza con los *hypernudges*. Yeung (2017) acuñó este término para describir intervenciones basadas en grandes volúmenes de datos que son dinámicas, continuas, ubicuas y adaptativas. Un *nudge* convencional introduce una modificación predeterminada y no adaptativa de la arquitectura de elección; por ejemplo, un aviso que invita al usuario a considerar la exactitud de una noticia antes de compartirla (Thaler y Sunstein, 2008; Pennycook et al., 2021). El *hypernudge*, en cambio, puede reconfigurar dinámicamente esa arquitectura a partir de los datos y del comportamiento del usuario (Yeung, 2017). La IA generativa añade la capaci-

dad de crear el mensaje a medida. La investigación ya muestra que los modelos pueden automatizar la persuasión psicológicamente personalizada a gran escala (Matz et al., 2024).

Esta arquitectura es ambivalente. Puede amplificar desinformación mediante microsegmentación opaca —la selección de grupos muy específicos mediante datos personales y criterios algorítmicos que los usuarios desconocen—, permitiendo adaptar los mensajes a sus vulnerabilidades, emociones y creencias previas, pero también puede hacer más eficaz la comunicación pública: adaptar explicaciones a la alfabetización sanitaria, responder dudas en distintos idiomas u ofrecer refutaciones personalizadas. El criterio no puede ser, por tanto, “IA sí” o “IA no”, sino quién define el objetivo, qué datos utiliza el sistema, qué grado de transparencia existe, cómo se valida la información, quién audita los resultados y qué capacidad conserva la persona para comprender y rechazar la influencia.

5. Qué puede hacer la economía del comportamiento

La economía del comportamiento no ofrece una vacuna única contra la desinformación, sino que aporta distintas intervenciones dirigidas a prevenir su aceptación, limitar su difusión o reducir sus consecuencias. Su eficacia aumenta cuando se combinan medidas sobre el individuo, el mensaje, la plataforma y la institución, y cuando se parte de un diagnóstico de las barreras concretas.

5.1. Nudges sobre la atención y el acto de compartir

Los *accuracy prompts* desplazan la atención hacia la exactitud antes de compartir. Una pregunta breve —“¿Considera que este contenido es fiable?”— o la valoración previa de un titular no relacionado pueden mejorar la calidad media de lo difundido sin

“Una política conductual contra la desinformación debería seguir un ciclo continuo: escuchar, diagnosticar, diseñar, probar y adaptar. La escucha social-con datos digitales y no digitales permite detectar preguntas, emociones, vacíos informativos y narrativas emergentes. No debe confundirse con vigilar opiniones ni con contar menciones sin contexto. La OMS recomienda que se realice con finalidad legítima, minimización de datos, protección de derechos y participación comunitaria (OMS, 2025b)”.

impedir la publicación. Un metaanálisis interno de veinte experimentos, con 26.863 participantes, encontró un efecto replicable y generalizable de este tipo de recordatorios, aunque con heterogeneidad entre contextos (Pennycook y Rand, 2022).

Las fricciones introducen una pausa en acciones que las plataformas han hecho casi automáticas: advertir que no se ha abierto el enlace, pedir una confirmación antes de reenviar muchas veces, limitar la difusión masiva o mostrar la fuente y la fecha con mayor prominencia. Una fricción pequeña puede ser apropiada cuando el coste social de la impulsividad es elevado. Debe, no obstante, aplicarse de forma proporcionada: demasiados avisos causan habituación y pueden obstaculizar también la circulación de información correcta.

Las etiquetas de verificación y las señales de credibilidad reducen la aceptación y difusión de contenidos marcados como falsos, pero plantean problemas de cobertura y efectos indirectos. Lo no etiquetado puede interpretarse erróneamente como verdadero, y una etiqueta sin explicación ofrece poca ayuda para evaluar mensajes futuros. También importa el encuadre de la corrección.

En diversos experimentos realizados con 9.512 personas de Argentina, Brasil, Chile y Colombia, presentar el mensaje como confirmación de un hecho verdadero generó más interacción y emociones más positivas que formular el contenido semánticamente equivalente como refutación de una falsedad (Arugete et al., 2024). La lección práctica es que conviene empezar por el hecho y no otorgar al bulo el protagonismo del mensaje.

5.2. Inoculación o *prebunking*

El *prebunking* o inoculación psicológica actúa preventivamente, antes de la exposición a la desinformación, advirtiendo sobre posibles intentos de manipulación y mostrando, como ciertas vacunas, versiones atenuadas de sus argumentos o de las técnicas empleadas, acompañadas de una refutación, con el fin de fortalecer la resistencia frente a futuros bulos (Lewandowsky y van der Linden, 2021). Podríamos decir, por tanto, que el *prebunking* inculca una versión debilitada de una estrategia manipuladora y

explica cómo reconocerla, a fin de robustecer nuestro sistema inmunocognitivo. En lugar de memorizar todos los bulos posibles, la persona aprende patrones transferibles: falsos dilemas, expertos impostados, apelaciones conspirativas, selección interesada de datos o uso de lenguaje emocional para eludir la evidencia. Juegos y vídeos breves de inoculación han mejorado la capacidad de identificar técnicas de manipulación y la calidad de las decisiones de compartir en diferentes culturas (Roozenbeek et al., 2022; van der Linden, 2022).

El *prebunking* es especialmente atractivo cuando se anticipa una narrativa -por ejemplo, dudas recurrentes antes de una campaña de vacunación- o cuando corregir después será costoso. Sus efectos pueden disminuir con el tiempo, por lo que conviene utilizar recordatorios o “dosis de refuerzo”. Debe evitarse, además, anunciar rumores muy marginales y darles una notoriedad que no tenían.

5.3. Corrección y *debunking*

Cuando el bulo ya circula, la corrección o *debunking* sigue siendo necesaria. Un desmentido eficaz debería: comenzar con el hecho correcto; advertir de forma clara, pero no estridente, sobre la falsedad; explicar por qué es engañosa; ofrecer una explicación causal alternativa; y repetir el hecho al final. La alternativa es crucial porque retirar una explicación sin reemplazarla deja un vacío mental que facilita volver al relato inicial. La evidencia disponible no respalda la idea de que corregir produzca habitualmente un “efecto rebote”; el riesgo existe en circunstancias específicas, pero no justifica la inacción (Ecker et al., 2022).

La fuente debe seleccionarse en función de la confianza del público, no solo de su jerarquía formal. Profesionales de atención primaria, farmacéuticos, asociaciones de pacientes, comunicadores científicos, líderes comunitarios o religiosos pueden ser más eficaces que una autoridad distante. El tono importa: ridiculizar al receptor amenaza su identidad y cierra la conversación. Corregir con empatía no significa equiparar evidencia y falsedad, sino reconocer la preocupación legítima, explicar la incertidumbre y preservar la relación.

5.4. Boosts y alfabetización

Los *boosts*, por su parte, no pretenden orientar directamente una decisión concreta, sino reforzar las competencias que permiten a las personas decidir por sí mismas. La alfabetización estadística y mediática, la lectura lateral o el aprendizaje de reglas sencillas para evaluar la credibilidad de las fuentes pueden aumentar la resistencia frente a futuras informaciones falsas, incluso cuando cambien sus contenidos o canales de difusión (Hertwig y Grüne-Yanoff, 2017; Kozyreva et al., 2020). En este ámbito se incluyen la alfabetización estadística, sanitaria, mediática y algorítmica. Algunas capacidades pueden enseñarse mediante reglas breves: identificar denominadores y riesgos absolutos; distinguir correlación y causalidad; comprobar autoría, fecha y fuente primaria; reconocer conflictos de interés; realizar lectura lateral, abriendo otras fuentes antes de profundizar en la página; o preguntar qué evidencia modificaría una afirmación.

Nudges y *boosts* no son sustitutos universales. Un recordatorio puede actuar de inmediato y a bajo coste, mientras que una competencia adquirida es más transferible y respetuosa con la agencia, pero exige tiempo, motivación y oportunidades de aprendizaje. La combinación es más prometedora: una fricción en el momento de compartir puede incorporar un microaprendizaje que explique qué verificar.

5.5. De intervenciones aisladas a una estrategia combinada

Una política conductual contra la desinformación debería seguir un ciclo continuo: escuchar, diagnosticar, diseñar, probar y adaptar. La escucha social -con datos digitales y no digitales- permite detectar preguntas, emociones, vacíos informativos y narrativas emergentes. No debe confundirse con vigilar opiniones ni con contar menciones sin contexto. La OMS recomienda que se realice con finalidad legítima, minimización de datos, pro-

tección de derechos y participación comunitaria (OMS, 2025b).

Sobre ese diagnóstico puede construirse una estrategia escalonada: información oficial accesible y oportuna; inoculación ante narrativas previsibles; recordatorios de exactitud y fricciones en el punto de difusión; correcciones rápidas mediante mensajeros confiables; y *boosts* que aumenten la resiliencia a largo plazo. La intervención debe evaluarse con resultados conductuales -apertura, contraste, difusión y decisiones de salud-, no solo con alcance o recuerdo. También han de medirse efectos adversos: fatiga, pérdida de confianza, desplazamiento hacia canales más opacos o aumento de desigualdades.

“La economía del comportamiento aporta una estrategia más amplia: rediseñar los contextos de exposición y difusión, inocular frente a técnicas manipuladoras, corregir con hechos y explicaciones alternativas, y fortalecer las competencias de la ciudadanía”.

6. Límites, heterogeneidad y ética

Los efectos medios de las intervenciones conductuales suelen ser modestos y heterogéneos. Esto no equivale a irrelevancia: una mejora pequeña puede evitar mucho daño si alcanza a millones de personas y su coste marginal es reducido. Pero tampoco debe magnificarse mediante porcentajes relativos o resultados de laboratorio difíciles de trasladar. DellaVigna y Linos (2022), al comparar estudios publicados con 126 ensayos de *nudges* que cubrían 23 millones de personas, estimaron un aumento medio de 1,4 puntos porcentuales en estos últimos,

frente a 8,7 puntos en la literatura académica; la potencia estadística, las características de las intervenciones y la publicación selectiva explicaban buena parte de la diferencia.

La evaluación debe informar cambios absolutos, intervalos de incertidumbre, duración, costes y distribución de efectos. Importa saber no solo si una intervención funciona en promedio, sino para quién, en qué canal y durante cuánto tiempo. Una etiqueta puede ayudar a usuarios con alta confianza en la fuente y resultar inútil para otros; un *boost* puede beneficiar más a quienes ya poseen tiempo y habilidades; una fricción puede reducir tanto false-

dades como información válida. La equidad ha de ser un resultado explícito, no una nota final.

La dimensión ética se intensifica con los *hypernudges*. La defensa clásica del *nudge* —que preserva formalmente la libertad de elección— es insuficiente cuando la arquitectura es invisible, personalizada y cambia de manera continua. Existe asimetría entre una plataforma que conoce el historial del usuario y una persona que desconoce por qué ve un mensaje. La personalización sanitaria puede implicar datos especialmente sensibles y explotar vulnerabilidades, miedos o estados emocionales.

Una gobernanza legítima debería exigir, al menos, finalidad pública explícita, evidencia de efectividad, proporcionalidad, transparencia comprensible, minimización de datos, auditoría independiente, mecanismos de reclamación y posibilidad real de optar por no participar. También debe separar la promoción de decisiones informadas de la imposición de una ortodoxia: gestionar desinformación no consiste en suprimir desacuerdos ni incertidumbres científicas. Cuanto mayor sea la capacidad de personalización y adaptación, más rigurosas deben ser la rendición de cuentas y la supervisión humana.

7. Implicaciones para los sistemas sanitarios latinoamericanos

América Latina reúne condiciones que hacen especialmente necesaria una respuesta institucional y, a la vez, desaconsejan copiar mecánicamente intervenciones diseñadas en otros entornos. El uso intensivo de redes y mensajería privada convive con brechas de conectividad, alfabetización y acceso sanitario; la comunicación circula en español, portugués y numerosas lenguas indígenas; y la con-

fianza en gobiernos, medios y sistemas de salud es desigual. Los canales cerrados dificultan la observación y la corrección, mientras que la proximidad de familiares y comunidades confiere gran credibilidad al mensaje reenviado.

La experiencia de la COVID-19 mostró la difusión regional de falsas curas, como el dióxido de cloro, y de narrativas antivacunas, junto con diferencias importantes entre países. El análisis de noticias falsas y mortalidad en seis países latinoamericanos subrayó la necesidad de sistemas de gestión de la infodemia adaptados a cada contexto (Nieves-Cuervo et al., 2021). No obstante, la agenda no debe quedar anclada en la pandemia; dengue, vacunación infantil, salud sexual y reproductiva, cáncer, salud mental, nutrición o resistencia antimicrobiana son ámbitos igualmente expuestos.

Cada ministerio o autoridad sanitaria debería contar con una función estable de gestión de desinformación, conectada con vigilancia epidemiológica, comunicación de riesgos, atención primaria y participación comunitaria. Sus tareas incluirían: escucha social ética; identificación rápida de vacíos de información; protocolos de respuesta; repositorios abiertos de mensajes y evidencias; coordinación con sociedades científicas y verificadores; y evaluación conductual. No es imprescindible crear grandes unidades centrales. Puede comenzarse con equipos pequeños y redes distribuidas que aprovechen capacidades existentes.

La intervención debe acercarse a los lugares donde se toman las decisiones. En mensajería privada pueden ofrecerse canales oficiales verificables, materiales reenviables y respuestas breves para profesionales y agentes comunitarios. En consultas y farmacias conviene disponer de guías para con-



**PAN.AS
SALUD**

**LA EMPRESA CON CASI 30 AÑOS
EN LA ASISTENCIA CLÍNICA Y PEDIÁTRICA
DE ALTA COMPLEJIDAD EN EL DOMICILIO.
CALIDAD Y PRESTIGIO EN SU CASA.**



Contacto: Sr. Martín Anro | Dto. Comercial - Prestaciones



11-2501-1104

versar sin confrontación. El contenido de las campañas debe probarse con sus destinatarios y adaptarse cultural y lingüísticamente. La cooperación regional permitiría compartir alertas, plantillas de inoculación e instrumentos de evaluación, manteniendo la adaptación local.

Por último, las restricciones presupuestarias refuerzan, no eliminan, la necesidad de evaluar. Los pilotos aleatorizados o cuasiexperimentales pueden comparar encuadres, mensajeros y momentos de envío con indicadores de conducta y costes. La meta no es maximizar clics, sino mejorar decisiones informadas, confianza y resultados de salud. Una intervención de efecto pequeño puede ser coste-efectiva si se integra en infraestructuras existentes; una campaña vistosa puede no serlo si solo alcanza a quienes ya estaban convencidos.

8. Conclusiones

La desinformación sanitaria prospera porque se adapta a nuestra atención limitada, ofrece relatos emocionalmente potentes y circula en entornos que premian la novedad, la identidad y la interacción. La IA generativa aumenta la escala y la capacidad

de personalización, y convierte la persuasión en un proceso dinámico. La respuesta no puede reducirse a publicar más datos ni a desmentir cada bulo cuando ya se ha difundido.

La economía del comportamiento aporta una estrategia más amplia: rediseñar los contextos de exposición y difusión, inocular frente a técnicas manipuladoras, corregir con hechos y explicaciones alternativas, y fortalecer las competencias de la ciudadanía. Estas herramientas deben combinarse con información oportuna, mensajeros confiables, escucha social y mejoras reales en la relación entre población e instituciones.

El criterio de éxito no es eliminar todo error -objetivo imposible e incompatible con una sociedad abierta-, sino reducir daños preservando la autonomía, el pluralismo y la confianza. Para ello se requiere experimentar, medir efectos absolutos y distributivos, abandonar intervenciones ineficaces y someter la personalización algorítmica a reglas exigentes. En los sistemas sanitarios latinoamericanos, construir esta capacidad antes de la próxima crisis puede ser tan importante como disponer de buenos mensajes durante ella. 

Referencias

- Abellán, J.M., y Jiménez-Gómez, D. (2020). Economía del comportamiento para mejorar estilos de vida y reducir factores de riesgo. *Gaceta Sanitaria*, 34(2), 197-199. <https://doi.org/10.1016/j.gaceta.2019.05.014>
- Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review*, 53(5), 941-973.
- Aruguete, N., Batista, F., Calvo, E., Guizzo-Altube, M., Scartascini, C., y Ventura, T. (2024). Framing fact-checks as a “confirmation” increases engagement with corrections of misinformation: A four-country study. *Scientific Reports*, 14, 3201. <https://doi.org/10.1038/s41598-024-53337-0>
- Carrasco-Farré, C. (2022). The fingerprints of misinformation: How deceptive content differs from reliable sources in terms of cognitive effort and appeal to emotions. *Humanities and Social Sciences Communications*, 9, 162. <https://doi.org/10.1057/s41599-022-01174-9>
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39(5), 752-766. <https://doi.org/10.1037/0022-3514.39.5.752>
- DellaVigna, S., y Linos, E. (2022). RCTs to scale: Comprehensive evidence from two nudge units. *Econometrica*, 90(1), 81-116. <https://doi.org/10.3982/ECTA18709>
- Eagly, A. H., y Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.
- Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K., y Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, 1, 13-29. <https://doi.org/10.1038/s44159-021-00006-y>
- Hasher, L., Goldstein, D., y Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior*, 16(1), 107-112. [https://doi.org/10.1016/S0022-5371\(77\)80012-1](https://doi.org/10.1016/S0022-5371(77)80012-1)
- Hertwig, R., y Grüne-Yanoff, T. (2017). Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science*, 12(6), 973-986. <https://doi.org/10.1177/1745691617702496>
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
- Kozyreva, A., Lewandowsky, S., y Hertwig, R. (2020). Citizens versus the Internet: Confronting digital challenges with

- cognitive tools. *Psychological Science in the Public Interest*, 21(3), 103-156. <https://doi.org/10.1177/1529100620946707>
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480-498. <https://doi.org/10.1037/0033-2909.108.3.480>
 - Lewandowsky, S., y van der Linden, S. (2021). Countering misinformation and fake news through inoculation and prebunking. *European Review of Social Psychology*, 32(2), 348-384. <https://doi.org/10.1080/10463283.2021.1876983>
 - Loomba, S., de Figueiredo, A., Piatek, S. J., de Graaf, K., y Larson, H. J. (2021). Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nature Human Behaviour*, 5, 337-348. <https://doi.org/10.1038/s41562-021-01056-1>
 - Matz, S. C., Teeny, J. D., Vaid, S. S., Peters, H., Harari, G. M., y Cerf, M. (2024). The potential of generative AI for personalized persuasion at scale. *Scientific Reports*, 14, 4692. <https://doi.org/10.1038/s41598-024-53755-0>
 - Mostajo-Radji, M. A. (2021). Pseudoscience in the times of crisis: How and why chlorine dioxide consumption became popular in Latin America during the COVID-19 pandemic. *Frontiers in Political Science*, 3, 621370. <https://doi.org/10.3389/fpos.2021.621370>
 - Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175-220. <https://doi.org/10.1037/1089-2680.2.2.175>
 - Nieves-Cuervo, G. M., Manrique-Hernández, E. F., Robledo-Colonia, A. F., y Grillo Ardila, E. K. (2021). Infodemic: Fake news and COVID-19 mortality trends in six Latin American countries. *Revista Panamericana de Salud Pública*, 45, e44. <https://doi.org/10.26633/RPSP.2021.44>
 - Organización Mundial de la Salud. (2024). Disinformation and public health. <https://www.who.int/news-room/questions-and-answers/item/disinformation-and-public-health>
 - Organización Mundial de la Salud. (2025a). Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models. <https://www.who.int/publications/i/item/9789240084759>
 - Organización Mundial de la Salud. (2025b). Social listening in infodemic management for public health emergencies: Guidance on ethical considerations. <https://www.who.int/publications/i/item/9789240108202>
 - Organización Panamericana de la Salud. (2020). La OPS no recomienda tomar productos que contengan dióxido de cloro, clorito de sodio, hipoclorito de sodio o derivados. <https://iris.paho.org/handle/10665.2/52484>
 - Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., y Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, 592, 590-595. <https://doi.org/10.1038/s41586-021-03344-2>
 - Pennycook, G., y Rand, D. G. (2022). Accuracy prompts are a replicable and generalizable approach for reducing the spread of misinformation. *Nature Communications*, 13, 2333. <https://doi.org/10.1038/s41467-022-30073-5>
 - Roozenbeek, J., van der Linden, S., Goldberg, B., Rathje, S., y Lewandowsky, S. (2022). Psychological inoculation improves resilience against misinformation on social media. *Science Advances*, 8(34), eabo6254. <https://doi.org/10.1126/sciadv.abo6254>
 - Rozin, P., y Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, 5(4), 296-320. https://doi.org/10.1207/S15327957PSPR0504_2
 - Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1), 99-118. <https://doi.org/10.2307/1884852>
 - Spitale, G., Biller-Andorno, N., y Germani, F. (2023). AI model GPT-3 (dis)informs us better than humans. *Science Advances*, 9(26), eadh1850. <https://doi.org/10.1126/sciadv.adh1850>
 - Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. En M. J. Metzger y A. J. Flanagin (Eds.), *Digital media, youth, and credibility* (pp. 73-100). MIT Press. <https://doi.org/10.1162/dmal.9780262625623.073>
 - Sunstein, C. R. (2014). *Why Nudge? The Politics of Libertarian Paternalism*. Reino Unido: Yale University Press.
 - Thaler, R. H., y Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.
 - Thaler, R. H., & Sunstein, C. R. (2021). *Nudge: The final edition*. Penguin Books.
 - Tversky, A., y Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2), 207-232. [https://doi.org/10.1016/0010-0285\(73\)90033-9](https://doi.org/10.1016/0010-0285(73)90033-9)
 - Tversky, A., y Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453-458. <https://doi.org/10.1126/science.7455683>
 - van der Linden, S. (2022). Misinformation: Susceptibility, spread, and interventions to immunize the public. *Nature Medicine*, 28, 460-467. <https://doi.org/10.1038/s41591-022-01713-6>
 - Vosoughi, S., Roy, D., y Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151. <https://doi.org/10.1126/science.aap9559>
 - Yeung, K. (2017). “Hypernudge”: Big Data as a mode of regulation by design. *Information, Communication & Society*, 20(1), 118-136. <https://doi.org/10.1080/1369118X.2016.1186713>