# Guía del alumno
# 3.9 · Efectos medidos y límites de la evidencia

Módulo 3 · Modelos, mecanismos y redes sociales | Capítulo 9 · Operaciones de influencia en redes sociales | L27 de 38

## G1 · Qué aprenderás y cómo trabajar

Una operación puede producir muchos mensajes y conseguir poca atención; puede captar atención sin cambiar actitudes, o favorecer una conducta sin alterar preferencias. La pregunta de esta lección es qué permite afirmar la evidencia en cada caso. El criterio de éxito del alumno consiste en elaborar un juicio que distinga medida, comparación, resultado e incertidumbre.

- RA1. Clasificar alcance, exposición, interacción, recuerdo, actitudes y conducta, indicando qué prueba y qué no prueba cada indicador.
- RA2. Explicar selección, confusión, causalidad y representatividad, comparando diseños antes/después, paneles y experimentos.
- RA3. Interpretar los estudios de Eady, Bond y el ensayo de plataformas de 2024 sin trasladar sus resultados a preguntas que no investigaron.
- RA4. Calcular una diferencia entre cambios y redactar una conclusión limitada por sus supuestos y datos.

## Itinerario de estudio

Primero estudia la presentación con la explicación del vídeo. Después trabaja G2–G6, resuelve G7 sin consultar G8 y contrasta tu razonamiento con las soluciones. Finalmente realiza el cuestionario limpio. Dedica aproximadamente 45–60 minutos al conjunto, además del vídeo de unos 8–9 minutos. Esta estimación es orientativa, no un tiempo de aprendizaje medido.

## Ubicación y límites

La 3.8 explicó formas de manipulación organizada. Aquí se aprovecha ese conocimiento para evaluar evidencia de efectos, sin repetir su catálogo de tácticas. La 3.10, «Identificación análisis crítico y resistencia», desarrollará indicadores de coordinación, atribución, falsos positivos y respuestas. Esta lección no cierra el capítulo 9 ni el módulo 3. No enseña a ejecutar campañas: enseña a analizar sus resultados.

Convención: F1–F8 identifican las fuentes comentadas del documento docente. Las páginas de PDF se cuentan desde la primera hoja del archivo, incluida la portada. Los casos de G7 son ficticios y sus cifras son exclusivamente didácticas.

# G2 · Qué mide realmente cada indicador

| Indicador | Medida posible | Límite decisivo |
|---|---|---|
| **Alcance** | Cuentas distintas registradas durante un periodo. | Cuentas no equivalen necesariamente a personas; una cuenta puede ser artificial. |
| **Exposición** | Oportunidad de encontrar un mensaje, o visualización registrada, según el método. | Una reconstrucción de cronología no demuestra atención. |
| **Interacción** | Clics, comentarios, compartidos o reacciones. | Puede expresar apoyo, crítica, burla o automatización. |
| **Recuerdo** | Reconocimiento o evocación del mensaje en una prueba. | Recordar no implica aceptar ni comprender. |
| **Actitud** | Evaluación favorable o desfavorable, medida con preguntas o escalas. | La respuesta puede variar por formulación, contexto o deseabilidad social. |
| **Conducta** | Acción observada o declarada: acudir, votar, comprar, etc. | La declaración y el registro independiente tienen limitaciones diferentes. |

Estas categorías no son peldaños que toda persona recorra en un orden obligatorio. Una conducta puede surgir por hábito, obligación o presión del contexto. El diagrama conceptual de Larson et al. distingue recepción, aceptación y efectos posteriores; no garantiza que cada emisión culmine en conducta (F4, pp. 55–57 / PDF 81–83).

## Unidad, denominador y periodo

Ejemplo ficticio: 12.000 visualizaciones entre 3.000 cuentas alcanzadas no significan 12.000 personas. El cociente es cuatro visualizaciones por cuenta en promedio; no significa que cada una haya visto exactamente cuatro veces. Especifica si se cuentan eventos, cuentas o personas, qué población forma el denominador y durante cuánto tiempo se registró.

## Fiabilidad y validez

Fiabilidad es consistencia de la medición en condiciones pertinentes; validez es el respaldo de la interpretación que se pretende hacer. Un contador puede registrar bien comentarios y medir mal aprobación, porque también cuenta rechazo. CESEDEN distingue ambas propiedades y advierte que muchos procesos cognitivos se infieren mediante indicadores observables (F5, p. 28 / PDF 28). Para evaluar aprobación hace falta una medida pertinente de actitud, no cambiar el nombre del contador.

# G3 · Cómo evaluar una atribución causal

## Selección y causas alternativas

La selección aparece cuando quienes reciben o buscan un contenido ya difieren de quienes no lo reciben. Por ejemplo, una preferencia previa puede explicar seguir a un emisor y apoyar después su propuesta. Esa preferencia es una causa común: confunde la asociación entre exposición y apoyo. La asociación sigue siendo un dato; su interpretación causal necesita más evidencia.

La cronología exige que la exposición preceda al cambio atribuido. Además, hay que considerar sucesos concurrentes, otras fuentes informativas y causalidad inversa: el interés puede generar exposición. Pregunta qué habría ocurrido sin la intervención. Ese contrafactual orienta la elección de una comparación; no puede observarse simultáneamente en la misma unidad.

| Diseño | Aporte | Objeción que permanece |
|---|---|---|
| **Antes/después** | Describe un cambio entre fechas. | Otros acontecimientos pudieron producirlo. |
| **Panel** | Sigue a las mismas personas y compara cambios. | Persisten confusores variables en el tiempo y exposición no aleatoria. |
| **Asignación aleatoria** | Hace comparables los grupos en expectativa; fortalece la atribución causal. | Pérdidas, contaminación entre grupos y medición defectuosa pueden comprometerla. |

## Validez interna y alcance externo

La primera pregunta es si el diseño sostiene la atribución dentro del estudio. La segunda es a quiénes, cuándo y dónde se puede extender. Una muestra grande reduce cierta incertidumbre estadística, pero no corrige por sí sola la exclusión de personas sin acceso a una plataforma. Asignar tratamientos al azar y seleccionar una muestra representativa son operaciones distintas.

## Comparar cambios

Si A pasa de 40 % a 52 % y B de 38 % a 49 %, los cambios son +12 y +11 puntos porcentuales. Su diferencia es +1 punto. No se resta solo el nivel final. Interpretar ese punto como efecto requiere que, sin la intervención, A y B hubieran seguido tendencias comparables; también exige descartar hechos diferenciales relevantes y comprobar mediciones equivalentes. Dos observaciones aisladas no prueban ese supuesto.

# G4 · Eady et al.: qué se pudo observar

## Pregunta, datos y contexto

Gregory Eady, Tom Paskhalis, Jan Zilinsky, Richard Bonneau, Jonathan Nagler y Joshua A. Tucker estudiaron la campaña estadounidense de 2016; publicaron el análisis en 2023. Relacionaron exposición potencial a cuentas de la Internet Research Agency en Twitter con actitudes, polarización percibida y preferencias o conducta electoral declaradas. El panel tenía 1.496 usuarios que consintieron vincular su cuenta y responder encuestas. Hubo tres olas durante la campaña y un contacto posterior para recoger voto declarado (F1, p. 9).

Reconstruyeron publicaciones de las cuentas seguidas por los participantes y detectaron las procedentes de operaciones extranjeras mediante listas publicadas por Twitter. «Exposición» significa oportunidad de encontrar publicaciones en esas cronologías; no certifica que el usuario las leyera. Tampoco puede validarse de forma independiente todo el procedimiento de identificación de cuentas de Twitter, cuya descripción pública era incompleta.

## Distribución y recepción

El 1 % de los participantes concentraba el 70 % de las exposiciones potenciales a cuentas rusas; el 10 % acumulaba el 98 %. La exposición era especialmente alta entre quienes se identificaban fuertemente como republicanos. En el último mes, la exposición procedente de medios nacionales y políticos estadounidenses era mucho mayor. Estos resultados describen distribución y contexto informativo; no acreditan aceptación del mensaje (F1, pp. 3–5, figuras 1–3).

## Resultado y explicaciones

Los análisis no detectaron una relación significativa y sustantiva entre exposición y los cambios de actitudes, polarización percibida o voto examinados. Los autores interpretan conjuntamente la concentración, el perfil previo de los receptores, el peso de otras fuentes y los resultados de los modelos como indicios concordantes de influencia individual limitada. Es una interpretación apoyada por esos análisis, no una demostración de que ninguna interferencia haya tenido efectos (F1, pp. 7–9).

El panel permite controlar características personales constantes al estudiar cambios dentro de cada participante. No convierte la exposición en aleatoria: incluso recibir un retuit incidental depende de a quién se ha decidido seguir. Las explicaciones alternativas no desaparecen solo por añadir controles estadísticos. La cifra agregada de circulación y la intención atribuida al emisor no sustituyen este examen de recepción y cambio.

# G5 · Interpretar resultados sin exagerarlos

## Magnitud e incertidumbre en Eady

La figura 6 de F1 compara mediante simulaciones la predicción bajo la exposición observada con la predicción sin exposición. Para el voto a Trump, la estimación central es −0,18 puntos porcentuales y el intervalo del 90 % va de −1,15 a +0,78. Se trata de una estimación modelada con datos observacionales, no de una diferencia experimental. Los autores llaman a esas bandas intervalos de predicción en la figura; no deben confundirse con los intervalos del 95 % de los coeficientes de la figura 5 (F1, pp. 7–8).

El intervalo incluye cero y valores de ambos signos. Esto limita la precisión de la conclusión; no significa que el efecto real sea exactamente cero ni que todos los valores internos sean igualmente probables. Un intervalo estrecho puede descartar magnitudes incompatibles con el modelo, pero no corrige una medición incompleta ni un sesgo de selección. La significación estadística tampoco indica por sí sola importancia práctica.

El estudio también aplicó pruebas de equivalencia: para 17 de 18 resultados de actitudes y polarización, rechazó relaciones de magnitud superior a 0,2 desviaciones estándar bajo sus modelos. Esa unidad expresa el tamaño respecto de la variabilidad de la medida. Esta prueba acota magnitudes; aporta más información que limitarse a no encontrar significación estadística, pero conserva las restricciones del diseño (F1, p. 6).

## Qué no resuelve este estudio

El resultado no cubre toda la interferencia extranjera, todas las plataformas ni todos los contenidos audiovisuales o canales posibles. Los autores plantean efectos de segundo orden, por ejemplo reacciones a la propia controversia sobre interferencia. Esa posibilidad es una hipótesis discutida, no un efecto causal medido por su panel. Tampoco se puede trasladar sin pruebas un promedio nacional a un estado concreto o a otro país y elección. F1, pp. 8–9.

## Contraste: Bond et al. (2012)

En Facebook, durante las elecciones de Estados Unidos de 2010, asignaron usuarios al azar a mensaje social, informativo o control. El social incorporaba imágenes de amigos que declaraban haber votado. Registraron participación mediante registros electorales: +0,39 puntos porcentuales frente al control. La comparación fortalece la inferencia causal en ese contexto; el enlace con registros limita la población observada. No estudia persuasión hacia un candidato ni operaciones encubiertas. F2, pp. 295–298 / PDF 1–4.

## Por qué pueden coexistir ambos hallazgos

Un resultado observacional sobre exposición a una operación y otro experimental sobre movilización no responden la misma pregunta. Cambian intervención, plataforma, población, fecha y variable de resultado. La coincidencia metodológica útil es exigir una medida concreta y una comparación explícita. La divergencia de resultados no autoriza a declarar universalmente eficaces o ineficaces las redes sociales.

# G6 · Ensayos de plataformas y fuentes de contexto

## La prueba de 2024

Gundars Bergmanis-Korāts y Tetiana Haiduchyk evaluaron la detección y retirada de manipulación comercial en Facebook, Instagram, YouTube, TikTok, VKontakte y X. En agosto y septiembre de 2024 utilizaron 44 publicaciones no políticas y compraron interacción artificial. A las cuatro semanas permanecía activa el 86 % de esa interacción. La unidad observada era actividad artificial, no ciudadanos, opiniones ni votos (F3, pp. 7 y 10 / PDF 7 y 10).

La intervención permite comprobar que, en esas condiciones, hubo actividad comprada que persistió. No es una asignación aleatoria de usuarios a mensajes para medir persuasión. Los autores reconocen limitaciones de escala; los proveedores, las plataformas y las formas de interacción no constituyen una muestra aleatoria de todas las manipulaciones. El resultado es histórico, no una comprobación del estado de las plataformas en noviembre de 2026. Su interpretación sobre vulnerabilidad institucional debe distinguirse del porcentaje observado (F3, pp. 25–26).

## Fuentes que documentan actividad, no una tasa de persuasión

Graphika y Stanford Internet Observatory describieron en Unheard Voice poca interacción en gran parte de los activos encubiertos analizados. Es información relevante sobre respuesta visible; no una medición directa de cambio de actitudes. Procede de datos de retiradas y análisis abierto, con límites de cobertura (F6, pp. 3–4 / PDF 4–5).

El inventario de Bradshaw, Bailey y Howard combina noticias, revisión documental, estudios de país y consulta experta. Advierte que hay operaciones no documentadas. Sirve para estudiar formas y presencia de actividad organizada, no para contar exhaustivamente operaciones ni calcular qué porcentaje de ciudadanos fue persuadido (F7, p. 6 / PDF 8). Helmus et al. advierten dificultades para medir impactos y reúnen indicios de distinta fuerza; no deben elevarse sus relatos o asociaciones a experimentos causales (F8, p. 14 / PDF 32).

## Ficha de lectura crítica reutilizable

- Pregunta y variable: ¿actividad, recepción, actitud o conducta?
- Población y tiempo: ¿quién entra, quién queda fuera y durante qué periodo?
- Comparación: ¿qué aproximación al contrafactual utiliza?
- Resultado: ¿magnitud, denominador e incertidumbre?
- Conclusión: ¿qué respalda y qué evidencia adicional haría falta?

Redacción modelo: «En [población y periodo], mediante [método], se observó [resultado]. Esto permite [conclusión delimitada]; no permite [salto inferencial], porque [limitación concreta]».

# G7 · Actividades: resuelve antes de consultar G8

## Actividad A · El informe del distrito (caso ficticio)

Una organización sostiene que su campaña informativa aumentó en doce puntos el apoyo a una propuesta. Se dispone únicamente de esta ficha. La campaña solo operó en A entre las dos mediciones. No hubo asignación aleatoria. Cada porcentaje procede de una nueva muestra de 500 adultos del distrito, seleccionada por el mismo procedimiento probabilístico. Se formuló la misma pregunta. No se han calculado intervalos ni se conoce la composición final o la no respuesta. Hubo una noticia nacional en ese periodo, pero no consta si afectó igual a ambos distritos.

| Apoyo a la propuesta | Antes | Después |
|---|---|---|
| **Distrito A** | 40 % | 52 % |
| **Distrito B** | 38 % | 49 % |

- A1. Identifica qué variable se midió y calcula los dos cambios y su diferencia.
- A2. Redacta una conclusión de 60–90 palabras. Distingue cambio observado y atribución causal.
- A3. Identifica tres comprobaciones que podrían fortalecer o debilitar esa atribución. No inventes sus resultados.

## Actividad B · Tres titulares (análisis documental)

Dispones de los resúmenes y límites de G4–G6; no hace falta consultar material externo para resolverla. Cada titular mezcla un dato real de esos estudios con una conclusión que debes evaluar.

- B1. «Eady demuestra que la interferencia extranjera nunca altera una elección».
- B2. «La actividad artificial persiste: el 86 % de los usuarios aceptó la propaganda».
- B3. «Pocas interacciones en las cuentas de Unheard Voice prueban que nadie resultó influido».

Para cada uno: identifica la medida real, señala el salto injustificado y redacta una alternativa de una o dos frases. Después explica qué dato adicional sería necesario para estudiar el efecto que el titular atribuye. Tiempo orientativo de ambas actividades: 15–20 minutos.

## Criterio de autocontrol

Tu respuesta debe conservar los resultados comprobados. Reconocer límites no consiste en declarar que nada puede conocerse ni en sustituir la evidencia por sospechas. Distingue lo observado, la interpretación y la información que falta.

# G8 · Soluciones razonadas

## Actividad A

A1. Se mide actitud declarada hacia una propuesta. A aumenta 12 puntos porcentuales y B aumenta 11. La diferencia de cambios es 1 punto: (52−40)−(49−38). No son porcentajes relativos. El incremento relativo de A sería 12/40=30 %, que responde a otra pregunta y no sustituye la comparación causal.

A2. Respuesta modelo: «El apoyo declarado aumentó doce puntos en A y once en B; la diferencia entre cambios es un punto. Estos datos no permiten atribuir doce puntos a la campaña. La atribución de un punto tampoco está demostrada: no hubo asignación aleatoria y desconocemos si las tendencias sin campaña y el impacto de otros hechos habrían sido comparables. También falta cuantificar la incertidumbre muestral y comprobar la composición y la no respuesta de las muestras».

A3. Son pertinentes: revisar tendencias anteriores; investigar noticias o acontecimientos específicos de cada distrito; comprobar exposición real y circulación entre distritos; examinar composición, no respuesta y aplicación de la encuesta; calcular incertidumbre con el diseño muestral. No basta el mismo tamaño muestral. No podemos declarar el punto estadísticamente significativo ni nulo sin ese análisis. La noticia nacional es una causa alternativa posible, no una explicación ya demostrada.

## Actividad B

B1. El panel relaciona exposición potencial en Twitter con cambios individuales; no evalúa toda interferencia, plataforma o elección. Alternativa: «Eady et al. no detectaron una relación relevante con los cambios estudiados en su panel de 2016; el diseño observacional y su cobertura limitan la conclusión». Para estudiar otros canales habría que medirlos y construir comparaciones adecuadas; no se presume su efecto.

B2. El 86 % corresponde a interacción artificial que persistía tras cuatro semanas. Alternativa: «El ensayo de 2024 encontró persistencia de actividad comprada en sus condiciones de prueba; no midió aceptación». Hacen falta medidas de actitud en destinatarios, antes y después o con grupos comparables, y datos de exposición. La retirada de interacciones por sí sola tampoco probaría desaparición de un efecto anterior.

B3. Se registró poca interacción en gran parte de los activos estudiados. Alternativa: «La respuesta visible fue limitada en esos datos; no se midió directamente toda influencia sobre los destinatarios». Para estudiar efectos se necesitan medidas de recepción y cambio con comparación apropiada. El silencio puede acompañar atención, indiferencia o rechazo; ninguna opción se impone sin datos.

## Cómo comprobar tu aprendizaje

Debes poder explicar por qué son insuficientes las conclusiones iniciales, corregirlas sin borrar los hallazgos y proponer datos que resuelvan una duda concreta. Si confundiste unidades, vuelve a G2; si atribuiste causalidad al cambio, a G3; si generalizaste un estudio, a G4–G6.

# G9 · Síntesis y consulta

El producto de una operación es lo que se difunde; el efecto es un cambio relevante cuya medición y atribución necesitan evidencia propia. Una investigación puede ser útil sin resolver toda la cadena. El análisis riguroso conserva esa utilidad y especifica sus fronteras.

| Pregunta de repaso | Dónde volver |
|---|---|
| **¿Qué representa el denominador de una cifra?** | G2 |
| **¿Quién habría cambiado aun sin el mensaje?** | G3 |
| **¿Se observó atención o solo oportunidad de exposición?** | G4 |
| **¿Qué magnitudes e interpretaciones permite el intervalo?** | G5 |
| **¿Se midió una persona o actividad de una plataforma?** | G6 |

## Lecturas principales

F1. Eady et al. (2023). Exposure to the Russian Internet Research Agency foreign influence campaign on Twitter in the 2016 US election and its relationship to attitudes and voting behavior. Nature Communications, 14, 62. Prioriza métodos y discusión, pp. 8–9; resultados, pp. 2–8. Paginación impresa y PDF coinciden.

### F1 · Artículo y materiales asociados

F2. Bond et al. (2012). A 61-million-person experiment in social influence and political mobilization. Nature, 489, 295–298. Fuente adicional. Ver diseño y figura 1.

### F2 · Publicación

F3. Bergmanis-Korāts y Haiduchyk (2024). Social Media Manipulation for Sale. NATO Strategic Communications Centre of Excellence. Método p. 7; resultado p. 10; límites pp. 25–26. Páginas impresas y PDF coinciden.

F4. Larson et al. (2009). Foundations of Effective Influence Operations. RAND, MG-654-OSD. Recepción, aceptación y conducta: pp. 55–57 / PDF 81–83; medición: pp. 65–67 / PDF 91–93.

### F4 · RAND

Las referencias complementarias F5–F8, los pasajes revisados y el alcance de cada evidencia están desarrollados en el documento docente de fuentes. Las lecturas amplían la lección; ninguna es indispensable para resolver las actividades o el cuestionario.

# G10 · Referencias complementarias

F5. CESEDEN / Centro Conjunto de Desarrollo de Conceptos (2020). Implicaciones del ámbito cognitivo en las Operaciones Militares. Documento de Trabajo 01/2020. Capítulo primero, «El problema de la medida», p. 28 / PDF 28. Útil para revisar fiabilidad, validez e indicadores de procesos no directamente observables.

F6. Graphika y Stanford Internet Observatory (2022). Unheard Voice: Evaluating Five Years of Pro-Western Covert Influence Operations. 24 de agosto. Síntesis y metodología, pp. 3–4 / PDF 4–5. Permite distinguir respuesta visible y efecto sobre actitudes, y comprender la selección de datos procedentes de retiradas.

F7. Bradshaw, Samantha; Bailey, Hannah; Howard, Philip N. (2021). Industrialized Disinformation: 2020 Global Inventory of Organised Social Media Manipulation. Working Paper 2021.1. Oxford, Project on Computational Propaganda. Metodología, p. 6 / PDF 8. Permite reconocer por qué un inventario de casos documentados no equivale a un censo ni a una medición de persuasión.

F8. Helmus, Todd C., y coautores (2018). Russian Social Media Influence: Understanding Russian Propaganda in Eastern Europe. RAND Corporation, RR-2237-OSD. Apartado «Impact», p. 14 / PDF 32. Reúne relatos y asociaciones de distinta fuerza; conviene identificar cuándo habla el informe y cuándo resume otra investigación.

F8 · Registro de la publicación

### Una pauta para profundizar

Escoge una de estas lecturas y completa la ficha de G6. Copia la localización, identifica la unidad observada y reformula la conclusión sin ampliar su alcance. Si la obra resume otra investigación, distingue esa referencia indirecta de una lectura del estudio original. No conviertas un propósito declarado en un resultado comprobado.