



COVID-19 en YouTube: Debates y polarización en la esfera digital

COVID-19 on YouTube: Debates and polarisation in the digital sphere

- Dr. Óscar Luengo. Profesor Titular, Departamento de Ciencia Política y de la Administración, Universidad de Granada (España) (ogluengo@ugr.es) (<https://orcid.org/0000-0001-6597-6459>)
- Dr. Javier García-Marín. Profesor Titular, Departamento de Ciencia Política y de la Administración, Universidad de Granada (España) (jgmarin@ugr.es) (<https://orcid.org/0000-0002-2766-0266>)
- Dra. Emíliana de-Blasio. Profesora Titular, Departamento de Ciencias Políticas, Universidad LUISS, Roma (Italia) (edeblasio@luiss.it) (<https://orcid.org/0000-0002-1505-9713>)

RESUMEN

Las redes sociales han transformado de forma muy significativa la forma en la que se produce el diálogo político, impulsando una configuración digital de la esfera pública. El presente artículo tiene como objetivo el análisis de la deliberación producida en las redes sociales, con un especial énfasis en la polarización. Tomando como referencia los comentarios observados en YouTube sobre la COVID-19 durante 2020 en España, Italia y Reino Unido, lo cual arroja una muestra de 111.808 comentarios, se aplicaron una serie de técnicas automáticas de análisis basadas en algoritmos, lo que supone una metodología cuantitativa novedosa en este ámbito de estudio. En línea con lo señalado por trabajos previos, la hipótesis que se plantea en este artículo es que el grado de polarización no se da con la misma intensidad en las esferas digitales de distintos casos. De esta manera, cabe esperar unos mayores registros de polarización en la esfera digital de los países del sur de Europa, adscritos a un modelo de pluralismo polarizado, que en países de otros modelos como el liberal. Los resultados confirman la hipótesis, verificando que no solo se observa mayor polarización en España e Italia que en Reino Unido, sino que, a nivel desagregado, los hallazgos apuntan a que la actividad más polarizante obtiene mayor aprobación en los países mediterráneos de nuestra muestra.

ABSTRACT

Social media has significantly transformed how political discussions and deliberations occur, mainly by providing a digital realm for the public sphere. This study aims to analyse the extent of polarised opinions across Spain, Italy and the United Kingdom regarding COVID-19 during 2020 within social media. To do this, we examined YouTube comments (n=111,808) using automatic analysis and machine-learning techniques based on algorithms. This methodological strategy denoted an innovative and unique quantitative approach for this field of study. In line with previous research, the hypothesis was that the degree of polarization does not crystallize in the same manner in different countries' digital spheres. Therefore, it could be said that higher levels of polarization occur amongst Southern European countries like Spain and Italy (both countries adhering to a polarised pluralism model), compared to other countries ascribing to the liberal model (the United Kingdom in our study), which provides evidence supporting previous research studies. The results confirmed the hypothesis that the polarization of digital deliberation between Spain and Italy is higher than in the United Kingdom. But, also, the findings based on more disaggregated analysis suggest that the most polarized attitudes are even rewarded by other users in Mediterranean countries.

PALABRAS CLAVE | KEYWORDS

Esfera pública, YouTube, ciberpolítica, aprendizaje profundo, polarización, COVID-19.
Public sphere, YouTube, cyberpolitics, deep learning, polarisation, COVID-19.

1. Introducción y estado de la cuestión

La aparición y el rápido desarrollo de las tecnologías digitales ha supuesto el comienzo de un cambio de paradigma que afecta tanto al ecosistema mediático como a la propia configuración de la esfera pública (López-García, 2005; Arias-Maldonado, 2016). Este concepto se ha definido por los teóricos como el espacio comunicativo en el que asuntos de interés general son considerados para su discusión por aquellos ciudadanos preocupados/afectados, cumpliendo una serie de criterios normativos ciertamente ambiguos (Dahlberg, 2004). Este proceso requiere el entendimiento mutuo de los participantes, así como un esfuerzo verdadero y honesto para alcanzar una conclusión consensuada (Serrano-Contreras et al., 2020).

El cambio de modelo, ya identificado por Jay Blumler (2018), fue puesto en relación con el concepto de crisis (Davis, 2019). La conexión parece aún más evidente en época de pandemia e incide sustancialmente en la transformación de la esfera pública. En este sentido, la teorización de los espacios públicos plantea una nueva perspectiva analítica basada en el concepto de horizontes públicos (Volkmer, 2014), que destaca la centralidad de unos medios en constante proceso de evolución, en los que se confirma el particular protagonismo de los digitales. El desarrollo de procesos de fragmentación extrema de la esfera pública contemporánea, acelerados por el entorno impuesto por la comunicación digital, han permitido acuñar una definición de transición, como la de «esfera post-pública» (Davis, 2019; Schlesinger, 2020; Sorice, 2020).

Las implicaciones de la expansión definitiva de los «social media» son de una envergadura enorme. La consolidación de distintos foros como Facebook, Twitter o YouTube, son un buen ejemplo de ello (Dogan & Smith, 2016). Las propias lógicas en la que se sustenta el funcionamiento de las redes sociales también han transformado el modo en que los fenómenos políticos son recibidos, percibidos y discutidos en la esfera pública (Fung et al., 2013), transformada ya en una esfera digital. No obstante, la literatura no da por hecho un efecto positivo en la profundización de los sistemas políticos, y podemos observar visiones y argumentos opuestos sobre la naturaleza, dinámica y perfil de esta alteración (Gozálvez-Pérez, 2011).

Por un lado, se ha señalado sistemáticamente al mundo digital como la esperanza para la promoción del compromiso cívico, al plantear un amplio catálogo de innovaciones democráticas que contribuyen a hacer más plural el discurso político, a facilitar una mayor implicación en lo público, a permitir que los ciudadanos vigilen y controlen al poder, a producir más información y proporcionar nuevos formatos de transmisión de los contenidos políticos y, eventualmente, a participar en la toma de decisiones. El denominado ciberoptimismo (Bruns, 2008) apunta a que las NTIC podrían servir para articular una conversación más distendida entre los ciudadanos, lo que no sucede en foros deliberativos muy formalizados, contribuyendo a aumentar su nodalidad (Margetts, 2009). A través de la conversación, los miembros de la sociedad clarifican sus propios puntos de vista, conocen las opiniones de otros interlocutores, y hallan los problemas a los que se enfrentan los ciudadanos (Stromer-Galley & Wichowski, 2011). De hecho, se considera que hablar de algunos temas de interés general con otros ciudadanos es una condición necesaria para la comprensión integral de la convivencia democrática y, por lo tanto, para darle un significado a la participación en la vida política (Rubio, 2000; Scheufele, 2001). En la medida en que la interacción es un componente vital de las sociedades democráticas, estos procesos podrían conducir hacia una deliberación pública más inclusiva y significativa (Bimber, 1998; Berry et al., 2010). En este sentido, las redes sociales muestran un gran potencial en la movilización y el empoderamiento de los ciudadanos, facilitando las opciones para interactuar entre ellos (y con sus representantes), ajenos totalmente a los canales de comunicación «mainstream» más institucionalizados y en manos de las grandes corporaciones mediáticas.

Por otro lado, un grupo de académicos escépticos más críticos (Fuchs, 2017) ha venido sosteniendo que, lejos del optimismo mencionado, las dinámicas que se observan apuntan a la fragmentación de esta esfera digital, y a la consolidación de burbujas de filtrado (filter bubbles) y de cámaras de resonancia (eco chambers), fenómenos cuyo escrutinio plantea una serie de retos metodológicos significativos y los cuales solo comenzamos a comprender recientemente (Pariser, 2011)¹. En la medida en que existen diferentes redes sociales y, dentro de estas, una variedad casi infinita de compartimentos diferenciados, los usuarios habituales acaban eligiendo a sus interlocutores. Se argumenta en este sentido que las redes constituyen

cámaras de resonancia donde solo escuchamos el eco de nuestra propia voz (Sunstein, 2008), por lo que la empatía social puede verse seriamente dañada al identificarse una supresión drástica de la exposición a la diversidad (Prior, 2007) y una clara homogeneidad ideológica (Valera-Orgaz, 2017). Asimismo, se entiende que el alto nivel de anonimato que se esconde en la red podría constituirse como un medio para exacerbar comportamientos comunicativos desinhibidos, avanzando en la dirección de un incremento de la discusión política maleducada, irrespetuosa, incívica o agresiva (Rowe, 2014). En este sentido, más que una deliberación racional e informada, las redes funcionan amplificando y modulando una atmósfera o estado de ánimo público a veces irreflexivo, manipulado y lleno de decibelios, que dificultan la reflexión serena, subrayándose la noción de resonancia afectiva (Fleig & von-Scheve, 2020). En consecuencia, uno de los riesgos más denunciados es el de la radicalización, la fragmentación y la polarización. Como apunta Sunstein (2008), dado que la fragmentación de la opinión pública puede producir un efecto reductor de la cohesión social, en las redes se podría dificultar el contacto entre distintos grupos de opinión y, de esa forma, profundizarse trazos de radicalización de las opiniones propias al no enfrentarse nunca a sus contrarias (Reese et al., 2007).

Al mismo tiempo, y quizá relacionado con lo anterior, la polarización, o polarización afectiva (Iyengar et al., 2019), se ha incorporado como uno de los principales rasgos de los fenómenos sociales y políticos de los últimos años, constituyéndose como un objeto de estudio de enorme importancia en el campo de la comunicación política. La polarización puede definirse como la distancia relativa entre dos puntos de vista políticos enfrentados. Sin embargo, se trata de un concepto en constante reajuste, aunque su acepción principal tiene que ver con el crecimiento del espacio entre polos provocado fundamentalmente por la influencia de las emociones y las creencias, más que por la razón y la evidencia (Mason, 2014; Olsson, 2013). Estas distinciones pueden conducir a posiciones extremas (Fletcher & Jenkins, 2019; Gidron et al., 2019).

El objetivo de esta propuesta reside en el estudio de la deliberación política desde la perspectiva de la polarización, que será rastreada de forma comparada en relación con la COVID-19, una cuestión que podría calificarse de lugar común (*common place*) al situarse en el ámbito de la salud pública. En concreto, este artículo explorará el nivel de polarización observado en la red YouTube durante 2020 en relación con este asunto en España, Italia y Reino Unido, países que representan modelos diferenciados en los estudios de comunicación política (Hallin & Mancini, 2004). Para ello, se analizarán los comentarios más relevantes en relación con este episodio empleando técnicas automáticas de análisis basadas en algoritmos, lo que supone una metodología cuantitativa novedosa en el ámbito de estudio que nos ocupa. La mayor parte de la investigación sobre deliberación política en los últimos años se ha centrado en las interacciones observadas en Twitter y Facebook (Bakshy et al., 2015; Conover et al., 2011; Gruzd & Roy, 2014; Jaidka et al., 2019; Oz et al., 2018). Detrás de este hecho no solo está la proyección de los liderazgos tradicionales, sino también una relativamente simple recopilación de información en términos prácticos a través de su API. YouTube, por otro lado, ha transitado desde una plataforma de repositorio de material audiovisual a un entorno que podría considerarse una red social de un tamaño varias veces mayor que Twitter (más de 2.000 millones de usuarios contra los 340 millones de Twitter) y que ofrece similares características de interacción con la inclusión de un sistema de recomendación, los «me gusta» (*likes*), junto con la administración de los datos (Allgaier, 2019).

2. Material y métodos

La hipótesis de investigación podría formularse en los siguientes términos: en los países adscritos a los modelos de pluralismo polarizado (España e Italia, en nuestro caso), cabe esperar una polarización mayor al analizar la deliberación política producida en la esfera digital en comparación con los países del modelo liberal (Reino Unido). Esta hipótesis encuentra sustento en las diferencias que se pueden hallar entre estos dos sistemas donde, además, en estos últimos años, la polarización se ha visto desafiada por las profundas transformaciones experimentadas como consecuencia tanto de los inesperados resultados electorales en algunos países, como por el impacto de algunas cuestiones que irrumpieron de forma abrupta en la opinión pública (ej., referéndum del Brexit). De ser positiva, se confirmarían los estudios que afirman que la polarización afectiva no es igual en todos los países (Boxell et al., 2020; Fletcher et al., 2020). Es decir,

que la polarización puede presentar una relación de dependencia con respecto a la sociedad analizada y no necesariamente con el medio de comunicación.

La muestra para el análisis se ha extraído de YouTube con los siguientes parámetros: los 50 vídeos con más visualizaciones en los sitios regionales de YouTube del Reino Unido, Italia y España («.co.uk», «.it» y «.es») y con las palabras clave de búsqueda «COVID» o «Coronavirus». Posteriormente, se extrajeron los comentarios de los usuarios para cada uno de los 150 vídeos, usando la propia API de la empresa, con un total de 111.808 (15.933 para España, 27.468 para Italia y 68.407 para el Reino Unido). Conjuntamente con los comentarios se descargaron otras variables, como el número de «likes» (me gusta), de respuestas, la identidad del autor, del vídeo, etc. Es importante notar que la extracción de información, que se realizó el 29 de diciembre de 2020, ha estado basada en la popularidad, por lo que no representa el periodo temporal de la crisis de la COVID-19, ni los temas de los vídeos han de ser forzosamente similares (más allá de tratar sobre la enfermedad). Hacerlo de otra forma alteraría los resultados de la investigación al eliminar la selección que hacen los ciudadanos al acercarse a esta red social.

Una vez adquirido el corpus, el texto se procesó siguiendo los estándares de este tipo de estudios: eliminación de símbolos atípicos, de espacios en blanco, tokenización, etc. (Meyer et al., 2008). Como el número de comentarios extraídos es masivo, no se ha considerado adecuada una técnica exploratoria cualitativa, tan útil en otros casos. Hemos juzgado necesaria la aplicación de una técnica automática de extracción y reducción de la información contenida en la muestra para su descripción. De esta forma podemos reducir la complejidad temática contenida en los casi 112.000 comentarios a unos pocos temas o ideas presentes en la mayoría de ellos (divididos por país). A este respecto, existen multitud de técnicas existentes, desde algoritmos supervisados (García-Marín & Calatrava, 2018) hasta los no supervisados (como los LDA y LSI). Como la muestra recoge textos muy variados, los algoritmos supervisados no parecen aconsejables, ya que el entrenamiento debería estar basado en todos los temas existentes en textos muy breves. Por ello, hemos decidido la aplicación de una técnica no supervisada, en concreto por un algoritmo LDA (Latent Dirichlet Allocation). Este algoritmo es una técnica enmarcada dentro del Procesamiento Natural del Lenguaje (NLP por sus siglas en inglés) que se basa en analizar las relaciones entre un conjunto de documentos y los términos que contienen (por lo que está basado en frecuencias) mediante la producción de conceptos relacionados con esos documentos y términos. El algoritmo asume que las palabras que tienen significados cercanos co-ocurrirán en textos similares. El resultado es una serie de temas presentes en cada uno de los documentos (una buena explicación de su funcionamiento y utilidad puede encontrarse ya en Letsche y Berry (1997) lo que muestra que es una técnica bastante consolidada).

El análisis de polarización es, sin embargo, más complicado. En primer lugar, se ha medido la polarización de cada unidad de análisis (comentario) a través del desarrollo de un índice propio. Esto se ha hecho así porque la medición de la polarización afectiva, definida como la identificación partidista, suele utilizar la encuesta como metodología de referencia (Druckman & Levendusky, 2019). Sin embargo, dicho método no puede aplicarse en los datos anónimos procedentes de las redes sociales (es decir, a textos). Para lograrlo se realizó una clasificación por medio de un análisis de sentimiento, que fue el encargado de calificar y clasificar la muestra. Como la muestra es multilingüe, se ha optado por utilizar un análisis idéntico para los tres idiomas (aunque adaptado a las lenguas a través de un diccionario específico para cada una de ellas) y no utilizar tres análisis diferenciados que pudieran sesgar los resultados. La herramienta seleccionada ha sido Orange3 (Demsar et al., 2013), basada en Python (que utiliza un diccionario multilingüe para más de 50 idiomas). No obstante, como ese tipo de análisis no aporta más información que una declaración sobre lo positivo o negativo de las palabras o frases, se elige una forma de medir la polarización pormenorizada (Serrano-Contreras et al., 2020). Para ello, hemos operacionalizado la polarización de un comentario como la distancia entre el análisis de sentimiento de ese texto y la mediana del agregado de los análisis de sentimiento de todos los comentarios del mismo país (España, Italia o Reino Unido), en número absoluto. De esta forma, obtenemos un registro que puede tomar cualquier valor entre 0 y 200, donde 0 muestra que no hay polarización (aunque pueda haber un sentimiento tanto positivo como negativo) y 200 muestra una polarización máxima. Es decir, definimos la polarización no como la expresión de la valencia, negativa o positiva, de una expresión, sino como la diferencia entre dichas

valencias. De tal forma que un comentario aislado no expresaría ningún tipo de polarización por sí mismo, pero sí si el contexto donde se produce es de signo contrario. Así, por ejemplo, un comentario negativo no es polarizante si la media de los mismos es igualmente negativa (como la condena a un acto criminal). Pero, en ese mismo ambiente, un comentario positivo sí que sería considerado como polarizante al situarse fuera de la media. Aunque la composición del índice pueda parecer sintética, los resultados ofrecidos son coherentes con la realidad percibida, como puede comprobarse más adelante. Naturalmente, el índice es lo suficientemente flexible como para permitir cambios en su composición: así, también hemos extraído la polarización por cada vídeo o por cada autor. Como la muestra es muy grande, del orden de las decenas de miles de comentarios, la polarización tenderá a cero, ya que existirán muchos elementos que se cancelarán o que, directamente, serán neutros. Por lo tanto, pequeños cambios en el número serán bastante significativos.

Por último, el índice de polarización ha servido como base para la realización de análisis estadísticos con el objetivo de modelizar su dependencia y exponer posibles diferencias entre estos países, que es la base de nuestra hipótesis.

3. Análisis y resultados

La Tabla 1, que resume la muestra capturada de YouTube, ha servido como fuente inicial para nuestra propuesta de investigación. Desde el punto de vista del análisis de frecuencias, se observa cómo los vídeos para el caso de España son los que menos comentarios registran, un poco más de la mitad de los comentarios para el caso italiano y menos de una cuarta parte de los comentarios para el británico. Esto podría ser consecuencia de unas cifras de población menores, aunque también pueden intervenir otros factores. Más allá de estas cuestiones, la magnitud de la muestra es suficientemente grande como para considerar los resultados del análisis estadístico significativos.

País	Videos (n)	Comentarios	Autores	Comentarios / Autor (DT)	«Likes»	«Likes» / Comentarios
España	50	15.933	7.291	2,18 (3,97)	49.102	3,08
Italia	50	27.468	9.578	2,87 (6,51)	83.972	4,69
Reino Unido	50	68.407	31.137	2,20 (5,05)	347.837	9,33

El resto de variables incluidas como posibles factores explicativos siguen un patrón similar: el español es la lengua con menos autores de comentarios y con menos «likes». Sin embargo, se pueden percibir algunas diferencias en la distribución de los casos. Por un lado, los autores de los comentarios en español intervinieron de forma menos activa (con una media de 2,18 comentarios), que los que lo hicieron en inglés e italiano (con valores respectivos de 2,20 y 2,87 comentarios). Por otro lado, los comentarios en el caso de los vídeos sobre España también recibieron menos promedio de «likes» (3,08) que en el de Italia (4,69) y el de Reino Unido (9,33). Naturalmente, este último hecho viene determinado por la cantidad de autores en cada idioma.

Si avanzamos más allá de un análisis descriptivo, podemos encontrar hallazgos relevantes en relación con el contenido de los comentarios. Como apuntábamos anteriormente, es complicado plantear una descripción de una muestra tan grande. Por este motivo decidimos emplear una técnica no supervisada, un LDA. Encontramos de especial interés que tras este proceso se han identificado diferencias sustanciales entre los tres conjuntos de texto. Los resultados de la aplicación del algoritmo se muestran en la Tabla 2.

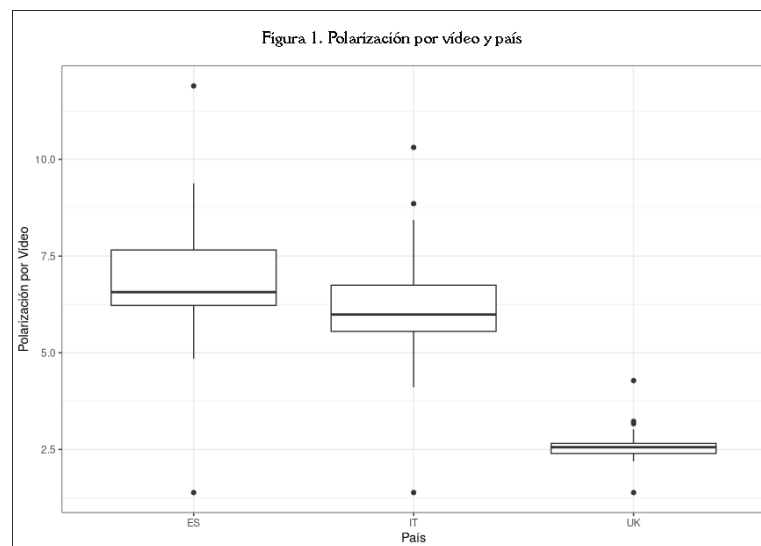
En los tres casos, las agrupaciones del algoritmo muestran un uso del lenguaje bastante peyorativo, aunque con diferencias sustanciales según la lengua. Por ejemplo, en el caso español el primer grupo está compuesto, en su mayoría, por insultos, no así en italiano o inglés, donde se aprecian palabras conectadas con los servicios públicos («scuola», «terapia») o el gobierno (government). El lenguaje despectivo está presente también en las muestras en inglés e italiano, al igual que las referencias a las instituciones en la muestra en español, pero no en el mismo orden (que es jerárquico según la importancia en la muestra), ni en la misma cantidad. Las alusiones a la política gubernamental también son frecuentes, pero mucho más para los comentarios en español e inglés, no así en italiano. Otro elemento interesante puede ser la aparición de elementos conspiranoicos y negacionistas en las muestras de inglés («Liverpool», «fake», «conspiracy») e italiano (negazionisti) pero no en la española.

España	Italia	Reino Unido
1: gracias, tonto, cojones, jajaja, ruina, jajajaja	1: grazie, grande, scuole, cervello, criminali, terapia	1: gracias, tonto, cojones, jajaja, ruina, jajajaja
2: si, mas, virus, gente, solo, España, estan, gobierno, mundo, pais	2: virus, solo, persone, covid,	2: si, mas, virus, gente, solo, España, estan, gobierno, mundo, pais
3: dictadura, puedes, malo, historia, madre, siendo, politico, dimision, nivel, iglesias	3: chiudere, morire, natale, negazionisti, informazione, chiudono, funziona	3: dictadura, puedes, malo, historia, madre, siendo, politico, dimision, nivel, iglesias
4: miguel, plan, jajajajajajaja, youtube, verano, gusto, corona, preveer	4: senso, colpa, bisogna, ridere, città, tampone	4: miguel, plan, jajajajajajaja, youtube, verano, gusto, corona, preveer
5: jose, che, don, ma, tontos	5: complimenti, azzimiei,	5: jose, che, don, ma, tontos

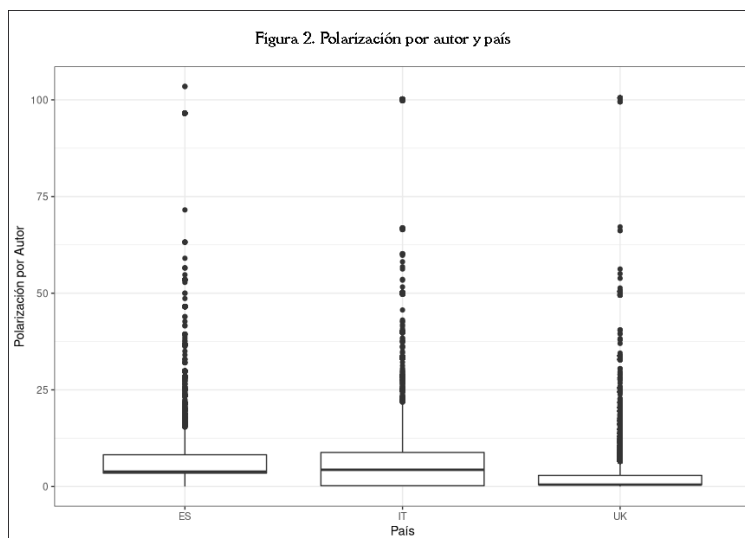
Para la descripción de la variable polarización, se ha decidido agrupar la muestra según idioma de los comentarios y vídeo o autor. Es decir, se ha medido la polarización media de los vídeos y de los autores de los comentarios, lo cual queda reflejado en la Tabla 3.

Polarización	Polarización	Polarización	Polarización	Polarización
Por vídeo				
Media (DT)	6,91 (1,50)	6,24 (1,32)	2,57 (0,37)	5,26 (2,24)
Mediana [min, max]	6,57 [1,39, 11,9]	5,99 [1,39, 10,3]	2,56 [1,39, 4,28]	5,74 [1,39, 11,9]
Por autor				
Media (DT)	7,78 (11,3)	7,08 (11,0)	2,68 (5,36)	4,33 (8,19)
Mediana [min, max]	3,85 [0,01, 103]	4,34 [0,06, 100]	0,48 [0,01,100]	1,70 [0,01, 103]

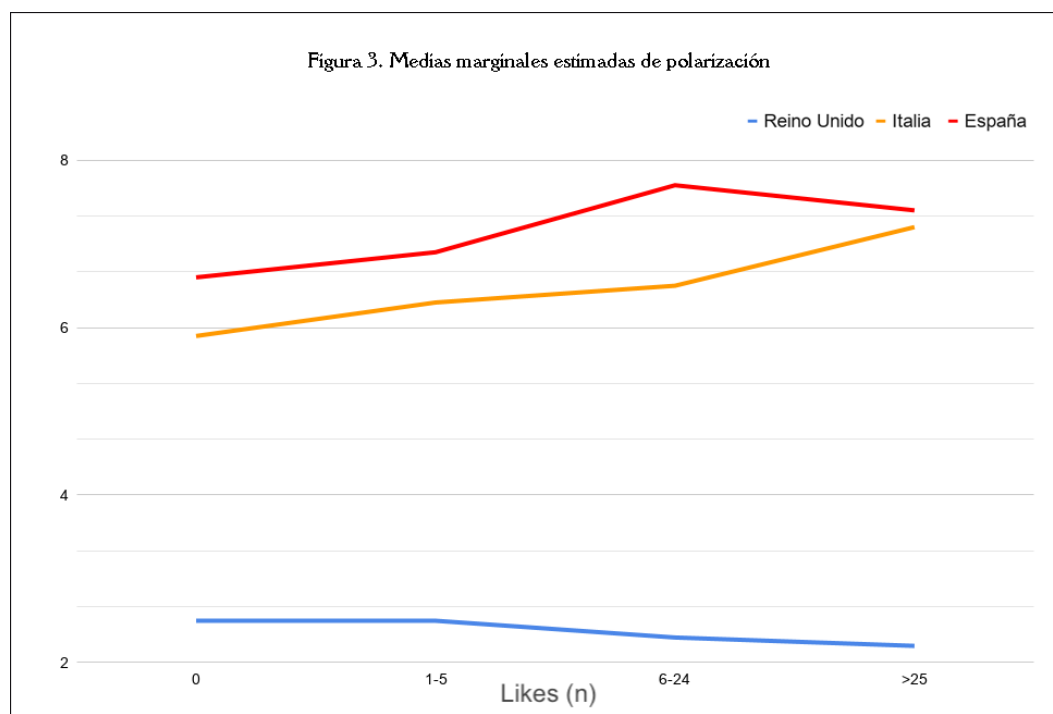
En el primer caso, son fácilmente observables las diferencias entre cada una de las lenguas analizadas: la polarización media es mucho mayor en el caso del español (6,91) e italiano (6,24) que en el caso del inglés (2,57). También resulta interesante observar cómo la desviación típica es mucho más pronunciada en el caso del español e italiano (1,50 y 1,32, respectivamente), que en el caso inglés (0,37). En todos los casos, la polarización media puede resultar excesivamente baja (después de todo el índice se movería teóricamente de 0 a 200). Sin embargo, hay que recordar que es una muestra muy grande compuesta, en su mayor parte, por términos no computables (preposiciones, artículos, emojis, etc.). Eso hace que, cuando se agrega todo el conjunto de datos, incluso por idioma, tienda a cero. De igual modo, podemos afirmar que, incluso obteniendo diferencias bastante pequeñas, siguen siendo visibles.



La Figura 1 representa de forma mucho más ilustrativa lo señalado anteriormente. Se observa muy claramente la obtención de dos conjuntos de datos diferentes: los comentarios en español e italiano, por un lado, y el caso inglés, por el otro. No solo llama la atención la agrupación de las medias segmentadas por caso (país), sino la dispersión tan amplia para los dos primeros casos, España e Italia.



En el caso del análisis de los comentarios desagregados por autor, los resultados muestran un patrón similar, como señala la Tabla 3: la media de la polarización por cada autor es mayor en los comentarios en español, seguidos por los comentarios en italiano y, a mucha distancia, los comentarios en inglés. Es igualmente interesante ver cómo la dispersión se comporta de una forma diferente: los mínimos y máximos en este caso son muy similares (de 0 a 100). Esto puede significar que ninguna muestra está exenta de autores altamente polarizantes («trolls», podríamos decir), aunque estos serían menos numerosos en el caso inglés que en el español e italiano. De nuevo, la Figura 2 muestra de forma más clara los resultados que los datos incluidos en las tablas: aunque la media inglesa sigue siendo inferior a la española e italiana, hay multitud de autores igualmente polarizantes, aunque tienden a concentrarse en la parte baja del gráfico.



Nota. Se han agrupado los «likes» por comentario en los cuatro grupos que pueden observarse en el eje x.

Naturalmente, las pruebas estadísticas más complejas reflejan una diferencia significativa² entre las variables «país» y «polarización» ($F=3,521$, $gl=2$; $p<0,05$). Las pruebas de Tamhane, T3 Dunnett y Games-Howell³ revelan que hay diferencias significativas entre los tres países y la polarización, aunque son más contundentes entre la muestra de Reino Unido y España (-4,22) e Italia (-3,54) que entre estos dos últimos (-0,67 Italia-España)⁴. Lo mismo sucede si agregamos la polarización por vídeo ($F=198,4$, $gl=2$, $p<0,05$; en este caso RU-ES -4,33, RU-IT -3,66 e IT-ES -0,67) o por autor de los comentarios ($F=1,969$, $gl=2$; $p<0,05$; RU-ES -5,10, RU-IT -4,40 e IT-ES -0,70).

Asimismo, al cruzar las variables «número de likes» y «país» en un modelo lineal univariado se han encontrado diferencias significativas que explican una buena parte de la variabilidad de la polarización (país*likes $F=9,259$, $gl=6$; $p<0,05$, $R^2=0,60$). Estos resultados son extremadamente interesantes porque muestran una pauta diferente para los tres países, explicando el comportamiento de la polarización en la muestra. De esta forma, podemos ver una divergencia entre la conducta de las personas que comentaron en los vídeos británicos con los que lo hicieron en los italianos o españoles, como con los datos presentados con anterioridad. Como se observa en la Figura 3, en el caso británico el número de «likes» desciende conjuntamente con la polarización. El caso italiano muestra un comportamiento totalmente contrario, siendo los comentarios más polarizantes los que tienen más «likes». El caso español es similar al italiano, con la divergencia de que los comentarios con más apoyo son algo menos polarizantes que los inmediatamente anteriores, pero no demasiado (y, en cualquier caso, siguen siendo más polarizantes que cualquier comentario en inglés o italiano).

Es importante apuntar que los «likes» son siempre posteriores a los comentarios, por lo que se podrían establecer relaciones causales. Aunque el modelo está lejos de explicar la totalidad de la muestra, consideramos muy significativo que en cada segmento (idioma) se premian o castigan los comentarios según su polarización. En esta línea, consideramos que las pautas culturales, más que las coyunturales, podrían explicar las diferencias. Por lo tanto, para la muestra estudiada, la variable «país» (idioma) es un buen predictor de la polarización y, dependiendo de la misma, el número de «likes» puede ser también una buena variable para explicar la polarización, ya sea de forma inversa (Reino Unido) o directa (Italia especialmente, pero también España). No se han encontrado, sin embargo, relaciones significativas entre el número de comentarios y la polarización por vídeo o autor, o entre el número de autores y la polarización, más allá de las meramente descriptivas, ya que la muestra está sobrerrepresentada a favor de los comentarios en los vídeos británicos.

4. Discusión y conclusiones

La investigación sobre la forma en la que se desarrolla la deliberación en la esfera digital se ha consolidado con firmeza como una prioridad en disciplinas como la ciencia política, las ciencias de la comunicación, la sociología o, incluso, las ciencias de la computación. Esta investigación propone una metodología innovadora para analizar la deliberación política en las redes sociales (YouTube), basada en el empleo de algoritmos para alcanzar una medida estandarizada de polarización.

Analizados los hallazgos, podemos confirmar que la hipótesis que vertebra esta propuesta queda verificada. En los países adscritos a los modelos de pluralismo polarizado (España e Italia), tratados como variable independiente, se observa una polarización mayor en la deliberación política en la esfera digital en comparación con los países del modelo liberal (Reino Unido). De igual forma, la evidencia mostrada específica que en los países del sur de Europa se premian los comportamientos polarizantes en la esfera digital, lo cual no solo no se produce en Reino Unido, sino que ocurre en sentido contrario. Siguiendo a Hallin y Mancini (2004), el primer grupo de países se caracteriza, entre otras cosas, por el denominado paralelismo político (militancia política de los medios). Por ello, interpretamos que este hallazgo se produce como consecuencia de la extrapolación del comportamiento de los medios de comunicación tradicionales a los espacios comunicativos digitales que se conforman en las redes sociales. Existiría pues un proceso de retroalimentación que contagiaría también las dinámicas comunicativas en la esfera digital, que podría consolidarse como una característica definitoria y confirmatoria de los perfiles que plantean los modelos referidos sobre los sistemas políticos y mediáticos, funcionando como elemento de extensión del paralelismo político, muy influenciado por la cultura política y mediática de cada uno de los sistemas, y

presente de forma más intensa en los modelos del sur de Europa. Esto ocurre, además, con un tema que puede considerarse como lugar común, en el que a priori se prevé menos polarización, aunque el papel de algunos actores muy activos en las redes, como los denominados negacionistas, convierten este punto en un tema muy interesante de discusión.

Uno de los resultados más llamativos de la presente investigación puede ayudar a dilucidar las diferencias entre los estudios teóricos y empíricos sobre la polarización y las redes sociales. Así, análisis como los de Spohr (2017), Parisier (2011) o Sunstein (2018) parecen indicar que la polarización podría ser un efecto sistémico del funcionamiento de la información en Internet y, más específicamente, de las redes sociales. Sin embargo, los análisis empíricos han cuestionado estos resultados (Bakhsy et al., 2015; Dougan & Smith, 2016; Boxell et al., 2017; Allcott & Gentzkow, 2017). Nuestro análisis puede aportar algo de información a este debate al encontrar diferencias regionales en el comportamiento de la polarización afectiva en redes sociales, aunque no sería la primera aportación en este sentido: Lee et al., (2014) ya apuntaron a que la polarización podría estar vinculada a las redes sociales, pero no con respecto a todos los temas. Algo similar muestra el estudio de Serrano-Contreras et al. (2020) al analizar la polarización con respecto a tres temas muy diferentes (elecciones en España, independencia de Cataluña y cambio climático). Y, sobre todo, parece apuntalar las tesis de Boxell et al. (2020), que han encontrado comportamientos muy diferentes entre países al analizar la polarización afectiva de 1980 a 2015 en 9 democracias avanzadas.

Las dinámicas de polarización descritas provocan un proceso de progresiva fragmentación de la esfera pública digital. La ruptura de la unidad de esta, por lo tanto, no es reemplazada por una segmentación plural de esferas públicas interconectadas (como parecía producirse en la fase inicial de las redes sociales, entre 1997 y 2002) sino por espacios públicos a menudo sin vínculos recíprocos o, en el mejor de los casos, con eslabones débiles y siempre bajo el condicionante de un fuerte proceso de polarización. La fragmentación de la esfera pública, acelerada por los ecosistemas de comunicación digital, produce una pulverización de experiencias y facilita la aparición de cámaras de resonancia, donde la acción de orientación operada por los algoritmos juega un papel no secundario en los mecanismos de información. Las conclusiones de nuestro trabajo están en consonancia con los resultados de investigaciones previas, que apuntan al crecimiento de las burbujas de filtrado, que juegan un papel de legitimación ideológica de las propias redes sociales y, más en general, de lo que recientemente se ha definido como «sociedad plataforma» (van-Dijck et al., 2018), y a los procesos de despolitización, desde las tendencias post-representativas de las democracias occidentales, agudizadas por las redes sociales, hasta la transformación sustancial de la esfera pública (Schlesinger, 2020; Sorice, 2020). Cabe señalar también que las redes sociales ayudan a formar islas de información que constituyen una caja de resonancia para generar vinculación comunicacional y saturación (Morlino & Sorice, 2021).

Los resultados de esta investigación plantean cuestiones adicionales, que podrían ser vetas de investigación futura. Por un lado, podrían incluirse más países en el análisis para comprobar si esta conexión entre polarización y fragmentación también es reconocible en otros países con sistemas político-institucionales diferentes. Por otro lado, la incorporación de otras plataformas como Twitter, Facebook o Instagram a esta misma lógica de análisis podría verificar si la estructura de polarización se da de la misma forma que en YouTube. En esta línea, comprobar el comportamiento de la deliberación sobre otros temas sería un interesante anclaje de comparación y contribuiría a una discusión más precisa de este objeto de estudio. Finalmente, la inclusión del control de otros aspectos como la estructura de las audiencias de los diferentes países, que podrían funcionar como variables intervinientes, podría suponer también un avance futuro.

Contribución de Autores

Idea, O.L., J.G.M.; Revisión de literatura (estado del arte), O.L., J.G.M., E.D.B.; Metodología, O.L., J.G.M.; Análisis de datos, O.L., J.G.M.; Resultados, O.L., J.G.M.; Discusión y conclusiones, O.L., J.G.M.; Redacción (borrador original), O.L., J.G.M.; Revisiones finales, O.L., J.G.M., E.D.B.; Diseño del Proyecto y patrocinios, O.L., J.G.M., E.D.B.

Notas

¹ El concepto de burbuja de filtrado se refiere a la dinámica de polarización determinada por el uso de algoritmos que determinan la creación de un sub-ecosistema digital. Es decir, es una propiedad definida por las posibilidades o por las cualidades estructurales de la red; las cámaras de eco, en cambio, se refieren a mecanismos homofílicos que, por tanto, tienen un carácter psicológico y juegan un papel importante en los procesos de aprendizaje posicional en la red.

² ANOVA con 1 factor, la prueba de Welch confirma que los datos no son iguales para los tres países, $p < 0,05$.

³ Se han utilizado estos estadísticos porque se asume que las varianzas no son iguales.

⁴ Las diferencias mostradas son del test de Tamhane pero son prácticamente idénticas en los otros dos test. También se mostrarán los resultados del test de Tamhane en el resto del artículo.

Apoyos

Grupo de Investigación en Ciencia Política y de la Administración (SEJ-113, Junta de Andalucía); Facultad de Ciencias Políticas y Sociología de la Universidad de Granada; Università LUISS (Roma).

Referencias

- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211-236. <https://doi.org/10.1257/jep.31.2.211>
- Allgaier, J. (2019). Science and environmental communication via online video: Strategically distorted communications on climate change and climate engineering on YouTube. *Frontiers in Communication*, 4(36), 1-18. <https://doi.org/10.3389/fcomm.2019.00036>
- Arias-Maldonado, M. (2016). La digitalización de la conversación pública: Redes sociales, afectividad política y democracia. *Revista de Estudios Políticos*, 173(173), 27-54. <https://doi.org/10.18042/cepc/rep.173.01>
- Bakshy, E., Messing, S., & Adamic, L.A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130-1132. <https://doi.org/10.1126/science.aaa1160>
- Berry, C., Kim, S., & Spigel, L. (2010). *Electronic elsewheres: Media technology and the experience of social space*. University of Minnesota Press. <https://bit.ly/33AbcKO>
- Bimber, B. (1998). The Internet and political transformation: Populism, community, and accelerated pluralism. *Polity*, 31(1), 133-160. <https://doi.org/10.2307/3235370>
- Blumler, J. (1995). The Crisis of Public Communication 1995-2017. *Javnost – The Public*, 25, 83-92. <https://doi.org/10.1080/13183222.2018.1418799>
- Boxell, L., Gentzkow, M., & Shapiro, J. (2020). *Cross-country trends in affective polarization*. National Bureau of Economic Research. <https://doi.org/10.3386/w26669>
- Boxell, L., Gentzkow, M., & Shapiro, J.M. (2017). Greater Internet use is not associated with faster growth in political polarization among US demographic groups. *Proceedings of the National Academy of Sciences*, 114(40), 10612-10617. <https://doi.org/10.1073/pnas.1706588114>
- Bruns, A. (2008). *Blogs, Wikipedia second life, and beyond: From production to produsage*. Peter Lang.
- Conover, M., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., & Flammini, A. (2011). Political polarization on Twitter. In N. Nicolov, & J. G. Shanahan (Eds.), *Proceedings of the International AAAI Conference on Web and Social Media* (pp. 89-96). The AAAI Press. <https://bit.ly/3bxioeB>
- Dahlberg, L. (2004). The Habermasian public sphere: A specification of the idealized conditions of democratic communication. *Studies in Social and Political Thought*, 10, 2-18. <https://bit.ly/2Nr1BBU>
- Davis, A. (2019). *Political communication: A new introduction for crisis times*. Polity. <https://bit.ly/3o85j17>
- Demsar, J., Curk, T., Erjavec, A., Gorup, C., Hocevar, T., Milutinovic, M., Mozina, M., Polajnar, M., Toplak, M., Staric, A., Stajdohar, M., Umek, L., Zagar, L., Zbontar, J., Zitnik, M., & Zupan, B. (2013). Orange: Data mining toolbox. *Python*, 14(1), 2349-2353. <https://bit.ly/3pMIPBR>
- Dougan, M., & Smith, A. (2016). *The political environment on social media*. Pew Research Center. <https://pewrsr.ch/2NyZVdh>
- Druckman, J., & Levendusky, M. (2019). What do we measure when we measure affective polarization? *Public Opinion Quarterly*, 83(1), 114-122. <https://doi.org/10.1093/poq/nfz003>
- Feinerer, I., Hornik, K., & Meyer, D. (2008). Text mining infrastructure. *Journal of Statistical Software*, 25(5), 1-54. <https://doi.org/10.18637/jss.v025.i05>
- Fleig, A., & Scheve, C. (2020). Introduction: Public spheres of resonance - Constellations of affect and language. In A. Fleig, & C. von Scheve (Eds.), *Public spheres of resonance. Constellations of affect and language* (pp. 1-16). Routledge. <https://doi.org/10.4324/9780429466533-1>
- Fletcher, R., Cornia, A., & Nielsen, R.K. (2020). How polarized are online and offline news audiences? A comparative analysis of twelve countries. *The International Journal of Press/Politics*, 25(2), 169-195. <https://doi.org/10.1177/1940161219892768>
- Fletcher, R., & Jenkins, J. (2019). *Polarisation and the news media in Europe*. European Parliamentary Research Service. <https://bit.ly/2ZKKpcQ>
- Fuchs, C. (2017). *Social media. A critical introduction*. Sage. <https://bit.ly/3tDtDsD>
- Fung, A., Gilman, H., & Shkabatur, J. (2013). Six models for the Internet and politics. *International Studies Review*, 15(1), 30-47. <https://doi.org/10.1111/misr.12028>
- García-Marín, J., & Calatrava, A. (2018). The use of supervised learning algorithms in political communication and media studies: Locating frames in the press. *Comunicación & Sociedad*, 31(3), 175-188. <https://doi.org/10.15581/003.31.3.175-188>

- Gidron, N., Adams, J., & Horne, W. (2019). How ideology, economics and institutions shape affective polarization in democratic polities. [Conference]. Annual Conference of the American Political Science Association, Washington DC, United States. <https://bit.ly/3aJkmJx>
- Gozálvez-Pérez, V. (2011). Education for democratic citizenship in a digital culture. [Educación para la ciudadanía democrática en la cultura digital]. *Comunicar*, 36, 131-138. <https://doi.org/10.3916/c36-2011-03-04>
- Gruzd, A., & Roy, J. (2014). Investigating political polarization on Twitter: A Canadian perspective. *Policy & Internet*, 6(1), 28-45. <https://doi.org/10.1002/1944-2866.poi354>
- Hallin, D., & Mancini, H. (2004). *Comparing media systems. Three models of media and politics*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511790867>
- Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S.J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, 22(1), 129-146. <https://doi.org/10.1146/annurev-polisci-051117-073034>
- Jaidka, K., Zhou, A., & Lelkes, Y. (2019). Brevity is the soul of Twitter: The constraint affordance and political discussion. *Journal of Communication*, 69(4), 345-372. <https://doi.org/10.1093/joc/jqz023>
- Lee, J.K., Choi, J., Kim, C., & Kim, Y. (2014). Social media, network heterogeneity, and opinion polarization. *Journal of Communication*, 64(4), 702-722. <https://doi.org/10.1111/jcom.12077>
- Letsche, T., & Berry, M. (1997). Large-scale information retrieval with latent semantic indexing. *Information Sciences*, 100(1-4), 105-137. [https://doi.org/10.1016/s0020-0255\(97\)00044-3](https://doi.org/10.1016/s0020-0255(97)00044-3)
- López-García, G. (2005). *Modelos de comunicación en Internet*. Tirant Lo Blanch.
- Margetts, H. (2009). Public management change and e-government: The emergence of digital-era governance. In A. Chadwick, & P. N. Howard (Eds.), *The Routledge Handbook of Internet Politics* (pp. 119-131). Routledge. <https://bit.ly/3bkxQel>
- Mason, L. (2014). 'I disrespectfully agree': The differential effects of partisan sorting on social and issue polarization. *American Journal of Political Science*, 59(1), 128-145. <https://doi.org/10.1111/ajps.12089>
- Morlino, L., & Soric, M. (2021). Quello che abbiamo appreso. In *L'illusione della scelta. Come si manipola l'opinione pubblica in Italia*. Luiss University Press.
- Olsson, E.J. (2013). A Bayesian simulation model of group deliberation and polarization. In F. Zenker (Ed.), *Bayesian argumentation: The practical side of probability* (pp. 113-133). Springer. https://doi.org/10.1007/978-94-007-5357-0_6
- Oz, M., Zheng, P., & Chen, G. (2018). Twitter versus Facebook: Comparing incivility, impoliteness, and deliberative attributes. *New Media & Society*, 20(9), 3400-3419. <https://doi.org/10.1177/1461444817749516>
- Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin. <https://doi.org/10.3139/9783446431164>
- Prior, M. (2007). *Post-broadcast democracy: How Media choice increases inequality in political involvement and polarizes elections*. Cambridge University Press. <https://doi.org/10.1017/cbo9781139878425>
- Reese, S., Rutigliano, L., Hyun, K., & Jeong, J. (2007). Mapping the blogosphere: Professional and citizen-based media in the global news arena. *Journalism*, 8(3), 235-261. <https://doi.org/10.1177/1464884907076459>
- Rowe, I. (2014). Incivility 2.0: A comparative analysis of incivility in online political discussion. *Information, Communication & Society*, 18(2), 121-138. <https://doi.org/10.1080/1369118x.2014.940365>
- Rubio, R. (2000). Internet en la participación política. *Revista de Estudios Políticos*, 19, 285-302. <https://bit.ly/3unSTVv>
- Scheufele, D.A. (2001). Democracy for some? How political talk both informs and polarizes the electorate. In R. P. Hart, & D. Shaw (Eds.), *Communication and U.S. elections: New agendas* (pp. 19-32). Rowman and Littlefield.
- Schlesinger, P. (2020). After the post-public sphere. *Media, Culture & Society*, 42(7-8), 1545-1563. <https://doi.org/10.1177/0163443720948003>
- Serrano-Contreras, I., García-Marín, J., & Luengo, O. (2020). Measuring online political dialogue: Does polarization trigger more deliberation? *Media and Communication*, 8, 63-72. <https://doi.org/10.17645/mac.v8i4.3149>
- Soric, M. (2020). La 'piattaformizzazione' della sfera pubblica. *Comunicazione Politica*, 3, 371-388. <https://doi.org/10.3270/98799>
- Spohr, D. (2017). Fake news and ideological polarization: Filter bubbles and selective exposure on social media. *Business Information Review*, 34(3), 150-160. <https://doi.org/10.1177/0266382117722446>
- Stromer-Galley, J., & Wichowski, A. (2011). Political discussion online. In M. C. Ess (Ed.), *The handbook of Internet studies* (pp. 168-187). Wiley-Blackwell. <https://doi.org/10.1002/9781444314861.ch8>
- Sunstein, C.R. (2008). *Republic.com 2.0*. Princeton University Press. <https://bit.ly/3vZ7R4b>
- Sunstein, C.R. (2018). *#Republic: Divided democracy in the age of social media*. Princeton University Press. <https://doi.org/10.2307/j.ctv8xnhtd>
- Valera-Orgaz, L. (2017). Comparing the democratic value of Facebook discussions across the profiles of Spanish political candidates during the 2011 General Election. *Revista Internacional de Sociología*, 75(1), 1-15. <https://doi.org/10.3989/ris.2017.75.1.15.119>
- van Dijck, J., de Waal, M., & Poell, T. (2018). *The platform society public values in a connective world*. Oxford University Press. <https://doi.org/10.1093/oso/9780190889760.001.0001>
- Volkmer, I. (2014). *The global public sphere: Public communication in the age of reflective interdependence*. Polity.