Appendices for Audible Reckoning: How Top Political Podcasts Spread Unsubstantiated and False Claims

Valerie Wirtschafter

February 8, 2023

Appendix 1: What Is an RSS Feed?

Developed in 1999 by Netscape and written in the web language XML, RSS feeds provide an RSS reader with instructions for what content resides on a given feed and how to read it.[1] The reader, which plugs into a series' content via a URL, is then automatically updated as new episodes are added to the feed. If data is removed from an RSS feed, it is no longer available through an RSS reader. While RSS feeds were once more commonly used for news headlines and blogs, they are now largely linked to the podcasting space, despite efforts to move away from the RSS architecture, which some executives in the space view as "outdated technology."[2] An example of a raw RSS file tied to the popular podcast "The Ben Shapiro Show" is included in figure A1.

Figure A1

Figure A1 shows what one episode of "The Ben Shapiro Show" looks like on an RSS feed.

In this example, the tags <item> and </item> bookend the episode; the feed also includes the episode title, description, publication date, episode number, author, summary, and duration. The MP3 file for this episode is also featured at the bottom, under <enclosure url = "URL TO EPISODE FILE" >. Each subsequent episode included in the RSS feed follows the same format.

After recording an episode, a podcaster uploads episode information to an RSS feed, hosted by a podcast hosting company for a small monthly fee. There are a wide variety of these hosting companies for setting up RSS feeds, but a few of the more commonly used ones are Libsyn, Podbean, and Blubrry. These feeds are then pushed to RSS readers like the Apple Podcast app. In the case of Apple

Podcasts, only approved podcasts that meet certain submission requirements can be listened to on the app.[3] Other RSS readers read any URL that a user passes to them automatically. The general pathway from audio recording to listener is detailed in figure A2.

Figure A2

CONTENT CREATION
Content recorded and edited by podcaster

STEP 1

CONTENT UPLOADED
Content uploaded to podcast RSS feed, hosted on an RSS site

STEP 2

CONTENT SHARED ON APP
RSS feed fed into an RSS reader (e.g., Apple Podcasts)

STEP 3

CONTENT ACCESSED BY LISTENERS
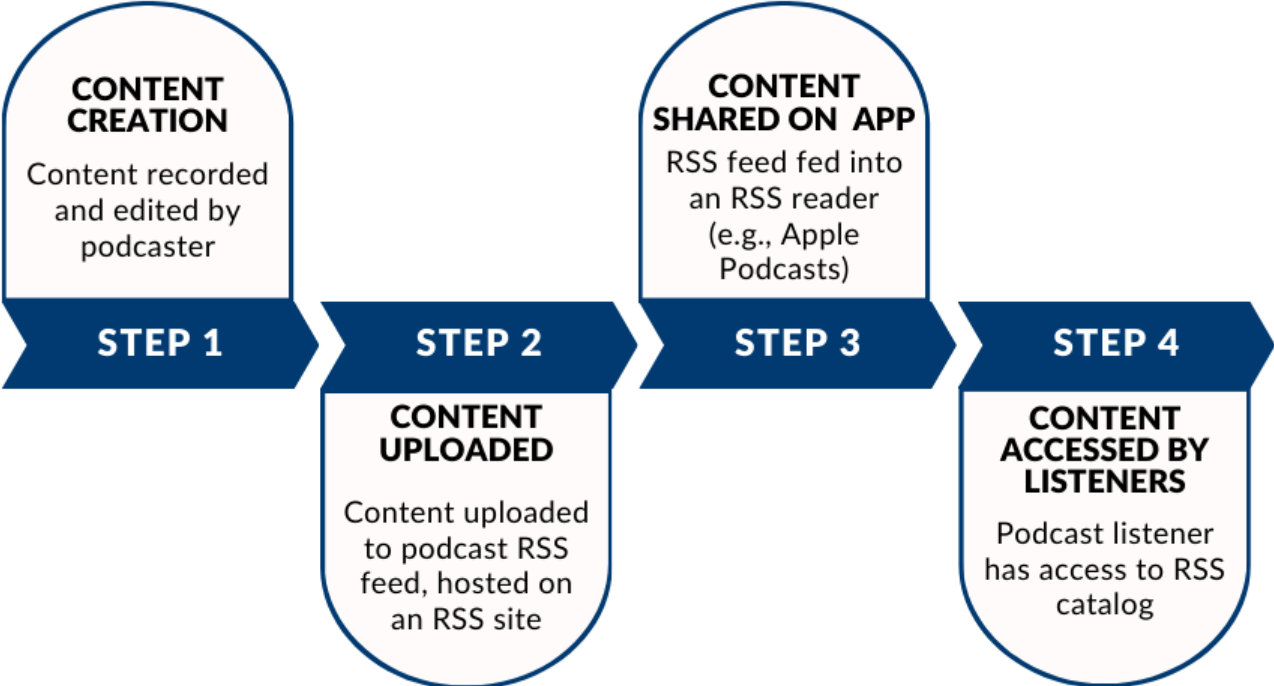Podcast listener has access to RSS catalog

STEP 4

Figure A2 details the podcast process from host recording the content to the audience listening to it.

Appendix 2: The Podcast Series in the Dataset

Show Name	Partisan Ideology	Total Episodes	Top 100 List?
"Bill O'Reilly's No Spin News and Analysis"	More Conservative	2192	
"The Sean Hannity Show"	More Conservative	2080	✓
"Bannon's War Room"	More Conservative	1750	✓
"The Glenn Beck Program"	More Conservative	1463	✓
"The Dan Bongino Show"	More Conservative	1307	✓
"The Daily Show With Trevor Noah: Ears Edition"	More Liberal	1221	✓
"Mark Levin Podcast"	More Conservative	1185	✓
"The Andrew Klavan Show"	More Conservative	1140	✓
"The Ben Shapiro Show"	More Conservative	1120	✓
"The Matt Walsh Show"	More Conservative	1118	✓
"Tim Pool Daily Show"	Unknown	1100	✓
"The Buck Sexton Show"	More Conservative	1077	✓
"The Michael Knowles Show"	More Conservative	1054	✓
"Uncle Joey's Joint With Joey Diaz"	Unknown	991	
"Conservative Review With Daniel Horowitz"	More Conservative	904	
"The News and Why It Matters"	More Conservative	881	
"The Clay Travis and Buck Sexton Show" (previously "The Rush Limbaugh Show")	More Conservative	869	✓
"Order of Man"	More Conservative	800	
"Late Night With Seth Meyers Podcast"	More Liberal	692	
"The Rubin Report"	More Conservative	600	
"Pod Save America"	More Liberal	484	✓
"The Adam Carolla Show"	More Liberal	479	✓
"The Axe Files with David Axelrod"	More Liberal	467	
"Political Gabfest"	More Liberal	450	
"Prager-U: Five-Minute Videos"	More Conservative	428	
"John Solomon Reports"	More Conservative	420	
"Stay Tuned With Preet"	More Liberal	415	
"Unashamed With Phil and Jase Robertson"	More Conservative	411	
"Louder With Crowder"	More Conservative	408	✓
"The Michael Savage Show"	More Conservative	395	
"Jocko Podcast"	Unknown	387	✓
"Pod Save the World"	More Liberal	363	
"Newt's World"	More Conservative	355	
"This Past Weekend"	Unknown	347	
"This Bulwark Podcast"	More Conservative	331	
"Pseudo-Intellectual With Lauren Chen"	More Conservative	313	
"The Charlie Kirk Show"	More Conservative	296	✓
"Get Off My Lawn Podcast With Gavin McInnes"	Unknown	270	
"The Fifth Column"	Moderate	262	
"Pod Save the People"	More Liberal	245	
"The Megyn Kelly Show"	More Conservative	245	✓
"The New Abnormal"	More Liberal	245	
"Federalist Radio Hour"	More Conservative	243	
"Lovett or Leave It"	More Liberal	232	✓
"Reasonable Doubt"	More Liberal	227	
"The DarkHorse Podcast"	More Conservative	224	✓
"Fireside Chat With Dennis Prager"	More Conservative	218	

Show Name	Partisan Ideology	Total Episodes	Top 100 List?
"Hold These Truths With Dan Crenshaw"	More Conservative	217	
"The All New Dennis Miller Option"	Unknown	215	
"Intercepted"	More Liberal	213	
"Rudy Giuliani's Common Sense"	More Conservative	207	✓
"Why Is This Happening? The Chris Hayes Podcast"	More Liberal	191	
"'Your Welcome' With Michael Malice"	Moderate	185	
"America First With Sebastian Gorka Podcast"	More Conservative	174	
"Congratulations With Chris D'Elia"	Unknown	174	
"Politicology"	Moderate	170	
"Firebrand With Matt Gaetz"	More Conservative	163	
"With Friends Like These"	More Liberal	159	
"The Jordan B. Peterson Podcast"	Unknown	148	✓
"Candace Owens"	More Conservative	144	✓
"The Laura Ingraham Show"	More Conservative	132	
"The MeidasTouch Podcast"	More Liberal	129	
"Majority 54"	More Liberal	118	
"The Ezra Klein Show"	More Liberal	116	✓
"The One With Greg Gutfeld"	More Conservative	112	
"Conversations with Coleman"	More Liberal	110	
"Verdict With Ted Cruz"	More Conservative	103	✓
"Real Time With Bill Maher"	More Liberal	100	✓
"The Lincoln Project"	More Conservative	95	✓
"WTF With Marc Maron Podcast"	More Liberal	87	
"Human Events Daily With Jack Posobiec"	More Conservative	80	
"The Candace Owens Show"	More Conservative	76	✓
"The National Pulse"	More Conservative	70	
"The Sarah Silverman Podcast"	More Liberal	67	✓
"The Portal"	Moderate	44	
"Honestly With Bari Weiss"	Moderate	39	
"Common Sense With Dan Carlin"	More Liberal	22	
"The Rachel Maddow Show"	More Liberal	22	✓
"All In With Chris Hayes"	More Liberal	17	

Appendix 3: Identifying Unsubstantiated or False Claims

Broad Analysis

To evaluate the spread of claims fact-checked as false across all podcast episodes, I first collected fact-checked claims from PolitFact and Snopes. In total, I collected 17,061 fact checks. Figure A3 includes a breakdown by fact check and rating.
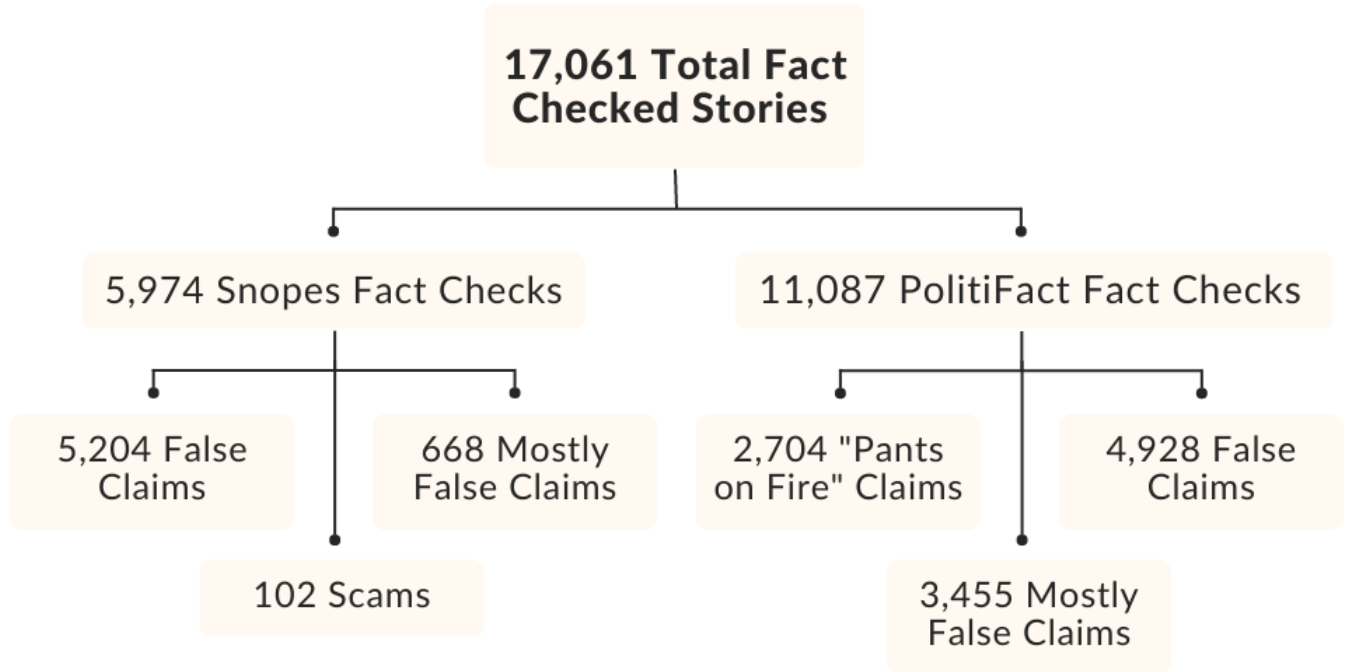
Figure A3: Overview of Fact-Checked Claims

Figure A3 provides an overview of the fact-checked claims searched across all podcasts episodes.

This analytical strategy involved sliding a window with a width 30% larger than the length of the fact check (for a 90-character-long fact check, the window allows for 120 characters) over the entire podcast episode and computed the cosine similarity between the podcast episode window and the fact-checked claim. This method then returned the best match for every fact check across all podcast episodes with a score detailing the quality of the match from 0 to 1 (with 1 in this case representing a better match and 0 representing a poor-quality match). I kept only matches with a cosine similarity of 0.5 or higher. I selected 0.5 as the cutoff because below this threshold, the quality of matches begins to deteriorate rapidly. By way of example, consider the fact- checked claim below in table A2 and the associated podcast episode window that triggered a match:

Table A2: Example of Cosine Similarity Results

Fact-Checked Claim	Podcast Episode Window
"you go to a gun show you can buy whatever you want and no background check"[4]	"go you go to a gun show you buy whatever you want no background check no thats completely a lie"[5]

In this case, the cosine similarly calculation returned a match of .781. However, a wide transcript window and/or manual review of this match revealed that the host was refuting the fact-checked claim. As a result, I did not code this podcast episode as sharing false information, at least as it pertained to this specific claim. Due to the buffer built into the podcast episode window, no claims would ever be a perfect match, but this buffer may have helped me identify slight word changes or offered additional context to speed up the manual review process.

Measuring False or Unsubstantiated Claims Tied to the 2020 Election

To evaluate the frequency with which political podcasters spread false or unsubstantiated election-related claims, first I subset the data to all episodes that aired between August 20, 2020, when Biden accepted the Democratic nomination, and the storming of the Capitol on January 6, 2021. This data range demarcated key moments in the 2020 election cycle and offered a picture of conversations on the potential for election fraud both before and after the election. Across the 79 series in the dataset, this analysis assessed the content from 3,916 episodes in total.

To code podcast episodes, I developed a dictionary of terms related to election fraud, drawing on popular lies, phrases, or statistics commonly cited during this period. This compendium included general terms like "election fraud" and more detailed terms like "missing usb cards," a reference to the widely disseminated and debunked narrative that 47 USB cards had gone missing from Pennsylvania voting machines.[6]

Episodes were evaluated (and flagged) to be sharing unsubstantiated or false election-related claims if the episode featured any form of editorializing about fraud narratives designed to cast doubt on the integrity of the election, as determined by at least two coders. This included highlighting the potential credibility of or fully endorsing at least one claim, spreading any claim (including reading quotes from someone else) without calling into question its veracity, or having a host feature a guest who repeated a false claim without being challenged.

Similar to prior coding exercises, two coders worked independently to evaluate whether a podcast episode was sharing unsubstantiated or false election-related claims, based on a match with the key term dictionary. If the two coders disagreed, a third coder adjudicated between their assessments to determine whether the episode shared unsubstantiated or false claims tied to the election. All terms used in this keyword dictionary can be found below.

<u>Election Fraud Terms</u>

The dictionary of terms related to election fraud drew on popular lies, phrases, or statistics commonly referenced before and after the 2020 election. These terms included: election fraud, software glitch, counting error, fraudulent biden elector, forensic electronic audit, forensic audit, voter integrity project, illegitimate president, election integrity, stop the steal, stop the steel, thrown out ballots, ditch along a wisconsin road, illegally flipped, missing usb cards, ballot stuffing, election irregularities, hugo chavez, indra, 800000 votes in pennsylvania, out of state license plates, 430 in the morning, venezuelan software, dominion company, 68 error rate, election hoax, surprise ballot dump, dominion systems, smartmatic, 450000 ballots, 3062 instances of voter fraud, stolen election, sharpies, dominion voting, arizona audit, voting machines, duplicate ballots, cyber ninjas, and rigged election.

Measuring False or Unsubstantiated Claims Tied to the Pandemic

Given the complexity of the evolving pandemic, this research took a conservative approach toward classifying false and unsubstantiated claims across popular political podcasters. Episodes were only flagged if they included verifiably false content that either contradicted scientific consensus (such as the notion that hydroxychloroquine usage could have saved thousands of lives), presented incorrect information (such as the claim that mask usage will kill you), or generalized broadly based on anecdotal evidence (such as the assertion that the vaccines cause spontaneous abortions).

Coding podcast episodes required a two-step filtering process to narrow down episodes in the dataset to those using terms or phrases commonly associated with unsubstantiated or false pandemic-related claims. The first step involved searching all the transcripts from episodes that aired between January 2020 and January 2022 for general terms tied to the pandemic (such as coronavirus, pandemic, and covid) to subset the data to episodes that at least mention the pandemic in passing. This eliminated episodes from the dataset that might have inadvertently matched a term in the COVID-19 dictionary but did not reference the pandemic (for example, terms like "fertility" or "bleach"). Then, using the dictionary of terms related to unsubstantiated or false pandemic-related claims, the second step involved filtering further to episodes that referenced one or more of these conspiracy theories or related terms.

As mentioned previously, the dictionary of COVID-19 terms drew on a wide variety of fact checks, articles, and reports detailing specific unsubstantiated or false claims. These terms were then searched for in all the transcripts. At least two coders evaluated any episode that was flagged as sharing one or more of these terms, and these coders also classified these claims in one of the six categories: (1) alternative treatments/prevention, (2) conspiracies, (3) government policy/response, (4) vaccine efficacy/side effects, (5) disease severity, and (6) testing.

In cases where guidance changed (on hydroxychloroquine usage, for instance) coders only flagged episodes as containing unsubstantiated or false pandemic-related claims after that change occurred. The coders did not flag any episodes as containing unsubstantiated or false pandemic-related claims for downplaying disease severity before March 13, 2020, when the Trump administration declared a national emergency due to COVID-19.[7] If an episode shared anecdotal information (such as a claim that my friend had fertility issues after the vaccine) coders did not flag that; however, coders did flag generalizable statements (such as the assertion that vaccines cause miscarriages). For conspiracy theories tied to the origins of the coronavirus pandemic, coders only focused on those that concerned intentional release (such as the view that the coronavirus was designed to remove Trump from office). Due to widespread controversy, coders did not code the "lab leak theory" as an unsubstantiated or false claim.[8] In addition, coders did not classify opinions as containing unsubstantiated or false claims. Take, for example, the mask debate during the coronavirus pandemic. Coders did not classify the claim "I think masks are child abuse" as an unsubstantiated or false claim, but coders *did* evaluate the claim "masks are harmful to your health" or "masks will kill you" as such. In addition, coders classified any absolute claim (such as the claim that COVID-19 has no effect on children/young adults) as an unsubstantiated or false claim but excluded qualified claims from this classification (such as the assertion that COVID-19 has less of an effect on children/young adults).

While this may seem like a conservative approach, these results reflected the dissemination of clear and obvious false claims as opposed to opinions or matters of political debate, however contentious they may be. Due to this conservative strategy, these results likely represent the lower numerical bound of unsubstantiated or false claims disseminated across podcasting.

COVID-19 Terms Dictionary

Terms used to identify episodes with unsubstantiated or false pandemic-related claims included: vaers, ivor mechtin, mass formation psychosis, robert malone, peter mccullough, hipaa violation, makeawish foundation, super viruses, mutant variant, plandemic, fort detrick, stolen from a canadian virus research lab, bioengineered, manufactured by the cia, population control, childrens health defense, forced vaccination, rare brain condition, vaccination information network, prion disease, vaccines cause miscarriages, secret ingredient, weed killer, 5g, tracking system, sterilization, roxana bruno, fertility, michael yeadon, robert kennedy, patricia finn, del bigtree, vaccine risk awareness movement, nanochip, microchip, andrew wakefield, suspicious death, millions against medical mandates, organic consumers association, national vaccine information center, lord fauci, rfk, mercola, charlene bollinger, ty bollinger, tenpenny, rizza islam, rashid buttar, erin elizabeth, sayer ji, kelly brogan, christiane northrup, ben tapper, kevin jenkins, meteor origin, casedemic, benny benassi, bat soup, betty white, bob saget, masks cause, mask are dangerous, vitamin d, vitamin c, common cold, bleach, hydroxychloroquine, ivermectin, miracle mineral supplements, silver infused, zithromax, zinc, more likely to be infected, more likely to be hospitalized, used to kill children, soft tissue cancer diagnos, not safe for pregnant women, more susceptible to hiv, gene therapy, athletes are collapsing, vaccines dont work, saline, spontaneous abortion, guillain barre syndrome, seasonal flu, spike proteins, pandemic of the unvaccinated, quarantine camp, populations growth, population control, undisclosed ingredients, dishonorable discharge, ida evacuees, illegal aliens who were covid positive, high risk people into camps, safer to be unvaccinated, miscarriage rate, never tested in humans, christian eriksen, vaccine related death, creating coronavirus variants, experimental jab, gmo shedding, dmx, pharmacovigilance tracking system, dying of natural causes, female sterilization, more suicide deaths than coronavirus death, improperly counted, doesn't affect children, children are immune, covid hoax, coronavirus hoax, stella immanuel, toxin inhalation, more illness than it prevents, deadliest vaccine, internment camp, strokes in pilots, parasite, withhold benefits, phony pandemic, undisclosed ingredient, population growth, toxic ingredient, luciferase, rna modifying, track vaccinated, bioweapon, inflated deaths, almost nothing, with not from covid, asymptomatic people cant, difference between the virus and the flu, education

camp, national database, court martial, unvaccinated veteran, shut down church, nuremberg code, force vaccinat, vaccines are more dangerous than, stillborn birth, vaccine is not approved

Coding Guidelines

Broadly speaking, the coders evaluated content as "containing unsubstantiated or false claims" if in their estimation it: (1) wholly endorsed a false or misleading claim; or (2) shared a claim in the context of a broader positive statement without directly refuting it. For example, in an episode titled "The Coming Rebellion," a guest on the Charlie Kirk Show reiterated the false claim that "Bank of America just pledged $1 billion dollars to Black Lives Matter" by stating that "Bank of America is giving away a billion dollars to Black Lives Matter, which is you know a fraudulent near terror organization."[9] This type of claim would fall under the category of "wholly endorses a false or misleading claim." In another instance, Sean Hannity shared a clip from a Trump event in an episode titled "The President Speaks the Truth" on January 9, 2019, in which the president falsely claimed, "at the request of Democrats [the border wall] will be a steel barrier rather than a concrete wall."[10] Following the excerpt from Trump's speech, Hannity commented, "this is a strong case, and I'm not worried about the president cause the president, as I described in detail earlier, you know has other options. He will declare a national emergency."[11] This instance would fall under the category "shares a claim in the context of a broader positive statement without directly refuting it."

Notes

[1] "RSS," Encyclopædia Britannica, https://www.britannica.com/technology/RSS.

[2] "Spotify 2022 Investor Day Transcript," Spotify, June 8, 2022, https://storage.googleapis.com/pr-newsroom-wp/1/2022/06/SPOTIFY-2022-INVESTOR-DAY-TRANSCRIPT.pdf.

[3] "Podcast RSS Feed Requirements," Apple Podcasts for Creators, https://podcasters.apple.com/support/823-podcast-requirements.

[4] Louis Jacobson and Amy Sherman, "In Gun Policy Address, Joe Biden Exaggerates about Background Checks at Gun Shows," PolitiFact, April 8, 2021, https://www.politifact.com/factchecks/2021/apr/08/joe-biden/gun-policy-address-joe-biden-was-wrong-about-backg/.

[5] Sean Hannity, "Meet Kathy Barnette," The Sean Hannity Show (Apple Podcasts), https://podcasts.apple.com/us/podcast/meet-kathy-barnette/id1112194905?i=1000516370462.

[6] Saranac Hale Spencer, "Nine election fraud claims, none credible," FactCheck.org, December 11, 2020, https://www.factcheck.org/2020/12/nine-election-fraud-claims-none-credible.

[7] "CDC Museum COVID-19 Timeline," Centers for Disease Control and Prevention, August 16, 2022, https://www.cdc.gov/museum/timeline/covid19.html.

[8] Amy Maxmen and Smriti Mallapaty, "The COVID Lab-Leak Hypothesis: What Scientists Do and Don't Know," Nature 594, no. 7863 (June 2021): 313–315, https://doi.org/10.1038/d41586-021-01529-3.

[9] Samantha Putterman, "No, Bank of America Didn't Pledge to Donate \$1 Billion to Black Lives Matter," PolitiFact, https://www.politifact.com/factchecks/2020/jul/23/facebook-posts/no-bank-america-didnt-pledge-donate-1-billion-blac/; Charlie Kirk, "The Charlie Kirk Show: Chris Bedford — The Coming Rebellion," YouTube, July 17, 2020, https://youtu.be/VoR8wM1tR3E?t=563.

[10] Manuela Tobias, "No, Democrats Didn't Ask for a Steel Barrier, as Donald Trump Said," PolitiFact, January 9, 2019, https://www.politifact.com/factchecks/2019/jan/09/donald-trump/no-democrats-didnt-ask-steel-wall-donald-trump/.

[11] Sean Hannity, "The President Speaks The Truth," The Sean Hannity Show (Apple Podcasts), https://podcasts.apple.com/nz/podcast/the-president-speaks-the-truth/id1112194905?i=1000427356794.