No. 23-411

IN THE

Supreme Court of the United States

VIVEK H. MURTHY, SURGEON GENERAL, *et al.*,

Petitioners,

v.

MISSOURI, *et al.*,

Respondents.

ON WRIT OF CERTIORARI TO THE UNITED STATES COURT OF APPEALS FOR THE FIFTH CIRCUIT

BRIEF OF *AMICI CURIAE* ELECTRONIC FRONTIER FOUNDATION AND CENTER FOR DEMOCRACY & TECHNOLOGY IN SUPPORT OF NEITHER PARTY

SAMIR JAIN
KATE RUANE
CENTER FOR DEMOCRACY & TECHNOLOGY
1401 K Street NW, Suite 200
Washington, DC 20005

DAVID GREENE
Counsel of Record
ELECTRONIC FRONTIER FOUNDATION
815 Eddy Street
San Francisco, CA 94109
(415) 436-9333
davidg@eff.org

Attorneys for Amici Curiae

325007

COUNSEL PRESS
(800) 274-3321 • (800) 359-6859

TABLE OF CONTENTS

Page

TABLE OF CONTENTS. i

TABLE OF CITED AUTHORITIES iv

STATEMENT OF INTEREST OF *AMICI CURIAE* 1

SUMMARY OF ARGUMENT 1

ARGUMENT 3

I. GOVERNMENT INVOLVEMENT IN CONTENT MODERATION INVOLVES COMPETING CONCERNS 3

A. Government Involvement in Content Moderation Raises Serious First Amendment and Broader Human Rights Concerns 3

B. Government May Productively and Appropriately Contribute to Content Moderation Practices 5

II. FIRST AMENDMENT DOCTRINE REFLECTS THESE COMPETING INTERESTS AND REQUIRES COURTS TO DISTINGUISH BETWEEN APPROPRIATE AND INAPPROPRIATE GOVERNMENT COMMUNICATIONS 7

Table of Contents

Page

A. *Bantam Books* Provides the Proper Framework For Assessing Improper Government Inducement of Private Censorship 8

B. The State Action Coercion Doctrine Should Be Read As Coextensive With the First Amendment Analysis Under *Bantam Books* 15

C. Close Cases Should Be Resolved Against the Government Since the Government Is Best Able to Model its Communications to Avoid Ambiguity 17

III. THIS COURT'S ANALYSIS MUST BE INFORMED BY A FULL UNDERSTANDING OF CONTENT MODERATION PROCESSES 18

A. Content Moderation Is an Historic and Widely Employed Practice With Varying Standards and Practices 19

B. Social Media Sites Commonly Consult With and Receive Advice and Feedback From Numerous External Sources, Including Governments 24

Table of Contents

Page

IV. EXAMINING SOME OF PETITIONERS' INVOLVEMENT IN CONTENT MODERATION 29

CONCLUSION 34

TABLE OF CITED AUTHORITIES

Page

Cases

Backpage.com v. Dart,
807 F.3d 229 (7th Cir. 2015) 4, 10, 11, 16, 33

Bantam Books v. Sullivan,
372 U.S. 58 (1963) 2, 4, 8, 9, 10, 15, 16, 33, 34

Blum v. Yaretsky,
457 U.S. 991 (1982) 15, 16

Bose Corp. v. Consumers Union,
466 U.S. 485 (1984) 2-3, 29

Carlin Comms., Inc. v.
Mtn. States Tel. & Tel. Co.,
827 F.2d 1291 (9th Cir. 1987) 16, 33

Hammerhead Enterprises, Inc. v. Brezenoff,
707 F.2d 33 (2d Cir. 1983), 464 U.S. 892 (1983) 13

Kennedy v. Warren,
66 F.4th 1199 (9th Cir. 2023) 11, 13, 14, 16

Miami Herald Publishing Co v. Tornillo,
418 U.S. 241 (1974) 4

Missouri v. Biden,
83 F.4th 350 (5th Cir. 2023) 34

Cited Authorities

Page

Missouri v. Biden,
No. 3:22-CV-01213, 2023 WL 4335270 (W.D. La. July 4, 2023) 27, 32, 33, 34

Morse v. Frederick,
551 U.S. 393 (2007). 6

Nat'l Rifle Ass'n of Am. v. Vullo,
49 F.4th 700 (2d Cir. 2022), ___S. Ct. ___, 2023 WL 7266997 (Nov. 3, 2023) 11

O'Handley v. Weber,
62 F.4th 1145 (9th Cir. 2023) 14

Okwedy v. Molinari,
333 F.3d 339 (2d Cir. 2003) 12

Penthouse Int'l, Ltd. v. Meese,
939 F.2d 1011 (D.C. Cir. 1991) 12

R.C. Maxwell Co. v. Borough of New Hope,
735 F.2d 85 (3d Cir. 1984) 12, 13

Rattner v. Netburn,
930 F.2d 204 (2d Cir. 1991) 12, 14

Rust v. Sullivan,
500 U.S. 173 (1991) 6

Speech First, Inc. v. Cartwright,
32 F.4th 1110 (11th Cir. 2022) 14

Page

VDARE Found. v. City of Colorado Springs, 11 F.4th 1151 (10th Cir. 2021), *cert denied*, 142 S. Ct. 1208 (2022). 12, 13, 16

Statutes

U.S. Const., Amend. I . . . . . 1, 2, 3, 4, 7, 8, 9, 15, 16, ...17, 28, 29, 32

20 U.S.C. § 7114(d)(6) 6

22 U.S.C. § 4411(b). 6

Other Authorities

About the YouTube Priority Flagger Program, YouTube, https://support.google.com/youtube/answer/7554338 (last visited December 15, 2023). 27

AJ Vicens, *Ahead of 2024 election, Meta worries about lack of information on top-tier nation-state covert operations*, Cyberscoop (Nov. 30, 2023), https://cyberscoop.com/ahead-of-2024-election-meta-worries-about-lack-of-information-on-top-tier-nationstate-covert-operations/ 6

Cezary Podkul, *Scammers Are Using Fake Job Ads to Steal People's Identities*, ProPublica (Oct. 26, 2021), https://www.propublica.org/article/scammers-are-using-fake-job-ads-to-stealpeoples-identities 22

Cited Authorities

Page

Community Guidelines, Instagram, https://help.instagram.com/477434105621119 (last visited December 15, 2023) 20

Community Guidelines, Peanut, https://www.peanut-app.io/community-guidelines (last visited Nov. 17, 2023). 23

Community Guidelines, Pinterest, https://policy.pinterest.com/en/community-guidelines (last visited December 15, 2023) 21

Community Standards Enforcement Report, Meta, Meta, https://transparency.fb.com/reports/community-standards-enforcement/ (last visited December 15, 2023) 26

Derek E. Bambauer, *Against Jawboning*, 100 Minn. L. Rev. 51 (2015) 14

Eric Goldman, *Content Moderation Remedies*, 28 Mich. Tech. L. Rev. 1 (2021). 20

Hannah Bloch-Wehba, *Automation in Moderation*, 53 Cornell Int'l L.J. 41 (2020). 19

Hannah Denham, *Another Fake Video of Pelosi Goes Viral on Facebook*, Wash. Post (Aug. 3, 2020), https://www.washingtonpost.com/technology/2020/08/03/nancy-pelosi-fakevideo-facebook/ 20

Page

Jillian C. York, *Silicon Values: The Future of Free Speech Under Surveillance Capitalism* (Verso 2021) 24

Kari Paul, *Naked Protesters Condemn Nipple Censorship at Facebook Headquarters*, The Guardian (June 3, 2019), https://www.theguardian.com/technology/2019/jun/03/facebook-nude-nippleprotest-wethenipple. 26

Kate Crawford & Tarleton Gillespie, *What Is a Flag For? Social Media Reporting Tools and the Vocabulary of Complaint*, 18 New Media & Soc'y 410 (2014) 25

Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 Harv. L. Rev. 1598 (2017) 20

Kevin Anderson, *YouTube Suspends Egyptian Blog Activist's Account*, The Guardian (Nov. 28, 2007), https://www.theguardian.com/news/blog/2007/nov/28/youtubesuspendsegyptianblog. 24

LinkedIn Professional Community Policies, LinkedIn, https://www.linkedin.com/legal/professional-community-policies (last visited December 15, 2023) 22

Meta Oversight Board, *UK Drill Music* (Jan. 2023), https://oversightboard.com/decision/IG-PT5WRTLW/ 28

Page

Meta, *How Stakeholder Engagement Helps Us Develop the Facebook Community Standards*, Meta (Jan. 18, 2023), https://transparency.fb.com/en-gb/policies/improving/stakeholders-helpus-develop-community-standards/. 25

Mike Masnick, *5th Circuit Cleans Up District Court's Silly Jawboning Ruling About the Biden Admin, Trims It Down To More Accurately Reflect The 1st Amendment*, Techdirt (Sep 11, 2023, 09:31 AM), https://www.techdirt.com/2023/09/11/5thcircuit-cleans-up-district-courts-silly-jawboning-ruling-aboutthe-biden-admin-trims-it-down-to-more-accurately-reflect-the-1st-amendment 29

Misinformation Policies, YouTube Help, https://support.google.com/ 21

Muhammed Imran Shafique, *How to Spot and Avoid LinkedIn Job Scams*, LinkedIn (Nov. 26, 2022), https://www.linkedin.com/pulse/how-spot-avoid-linkedin-job-scams-imran-shafique. 22

Naomi Appelman & Paddy Leerssen, *On 'Trusted' Flaggers*, 24 Yale J.L. & Tech. 452 (2022). 27

NHTSA, *Buzzed Driving is Drunk Driving*, https://www.nhtsa.gov/campaign/buzzed-driving (last visited July 26, 2023). 7

Page

Rebecca Kern and Maggie Miller, *Meta says US stopped sharing foreign threat intel on elections*, Politico (Nov. 30, 2023), https://subscriber.politicopro.com/article/2023/11/meta-says-usstopped-sharing-foreign-threat-intel-on-elections-00129181 34

Reducing Hate And Disinformation Online, Change the Terms, https://www.changetheterms.org (last visited December 15, 2023) 21

Robyn Caplan, *Networked Governance*, 24 Yale J.L. & Tech. 541 (2022) 24

Rules Enforcement, Twitter, https://transparency.twitter.com/en/reports/rules-enforcement.html#2021-jul-dec. 26

Safety Partners, TikTok, https://www.tiktok.com/safety/en-us/safety-partners/ (last visited December 15, 2023) 25

The Santa Clara Principles, https://santaclaraprinciples.org 3

Scams and Fraud Content, LinkedIn, https://www.linkedin.com/help/linkedin/answer/a1338803 (last visited December 15, 2023) 22

Page

Sensitive Media Policy, Twitter (March 2023), https://help.twitter.com/en/rules-and-policies/media-policy (permitting "consensually produced adult nudity") 20

Seny Kamara *et al.*, *Outside Looking In: Approaches to Content Moderation in End-to-End Encrypted Systems*, Ctr. for Democracy & Tech. (2021), https://cdt.org/wp-content/uploads/2021/08/CDT-Outside-Looking-In-Approaches-to-Content-Moderation-in-End-to-End-Encrypted-Systems-updated-20220113.pdf. 19

Seth F. Kreimer, *Censorship by Proxy: The First Amendment, Internet Intermediaries, and the Problem of the Weakest Link*, 155 U. Pa. L. Rev. 11 (2006) 14

The YouTube Team, *Our Ongoing Work to Tackle Hate*, YouTube (June 5, 2019), https://blog.youtube/news-and-events/ourongoing-work-to-tackle-hate/ 21

YouTube Community Guidelines Enforcement, Google, https://transparencyreport.google.com/youtube-policy/removals (last visited December 15, 2023) 26, 27

STATEMENT OF INTEREST OF AMICI CURIAE[1]

Electronic Frontier Foundation is a member-supported, nonprofit civil liberties organization that has worked for over 30 years to protect free speech, privacy, security, and innovation in the digital world on behalf of its nearly 30,000 members. EFF frequently files briefs in cases addressing online intermediary content moderation, and studies and writes extensively on the issue.

Center for Democracy & Technology is a non-profit public interest organization that for more than 25 years has worked to ensure that the constitutional and democratic values of free expression and privacy are protected in the digital age. CDT regularly advocates in support of First Amendment rights on the Internet and other protections for online speech, including for user-generated content. CDT also has decades of experience advising and advocating to online services about adopting trust and safety practices that best serve their users.

SUMMARY OF ARGUMENT

Government co-option of the content moderation systems of social media companies is a serious threat to freedom of speech, particularly for users of social media services with no say in whether their speech is silenced; but there are clearly times when it is permissible, appropriate,

1. No counsel for a party authored this brief in whole or in part, and no such counsel or party made a monetary contribution intended to fund the preparation or submission of this brief. No person other than *amici curiae*, or their counsel, made a monetary contribution intended to fund its preparation or submission.

and even good public policy for government agencies and officials to noncoercively inform, communicate with, and even attempt to persuade social media companies about the user speech they publish on their sites.

The applicable First Amendment test, from *Bantam Books v. Sullivan*, 372 U.S. 58, 71 (1963), balances these competing interests. The First Amendment forbids the government from intimidating or coercing a private entity to censor, whether that pressure is direct or informal. *Id.* But the test also recognizes that not every communication to an intermediary about users' speech is unconstitutional. *See id.* at 71-72 ("We do not hold that law enforcement officers must renounce all informal contacts . . . .").

The distinction between proper and improper government communication influencing publishing decisions is often fuzzy. But the government is best placed to provide clarity, by ensuring that its communications with intermediaries neither directly nor informally serve as regulatory instruments. This Court should clarify the analysis to assist all stakeholders and ensure essential communications can resume with proper safeguards for speech in place.

Neither the district court nor the Fifth Circuit adequately distinguished between improper and proper communications nor provided adequate guidance to either the government or anyone seeking to hold the government to its proscriptions.

This Court must independently review the record and make the searching distinctions the lower courts did not. *See Bose Corp. v. Consumers Union*, 466 U.S. 485,

499-500 (1984). This amicus brief aims to provide useful information about the competing interests involved and the environment of social media content moderation in which they must be applied.

ARGUMENT

I. GOVERNMENT INVOLVEMENT IN CONTENT MODERATION INVOLVES COMPETING CONCERNS

A. Government Involvement in Content Moderation Raises Serious First Amendment and Broader Human Rights Concerns

Government involvement in private companies' content moderation processes raises serious human rights concerns. The issue is highlighted in the Santa Clara Principles, of which *amici curiae* are among the co-authors, which aim to foster transparency and accountability in the moderation of user-generated content. The Principles specifically scrutinize "State Involvement in Content Moderation" and affirm that "state actors must not exploit or manipulate companies' content moderation systems to censor dissenters, political opponents, social movements, or any person."[2] "Special concerns are raised by demands and requests from state actors (including government bodies, regulatory authorities, law enforcement agencies and courts) for the removal of content or the suspension of accounts."[3]

---

2. The Santa Clara Principles, https://santaclaraprinciples.org (last visited July 26, 2023).

3. *Id.*

Government involvement in content moderation implicates both users' and companies' First Amendment rights. Users have an obvious First Amendment right against government censorship of their constitutionally protected speech online, even when such censorship is accomplished subtly. *Bantam Books*, 372 U.S. at 69–71. And online publishers have a First Amendment right to decide what speech to publish and how to publish it. *See Miami Herald Publishing Co v. Tornillo*, 418 U.S. 241 (1974).

Courts must be especially protective of users' rights because users cannot depend on the intermediaries, on whom they rely to distribute their speech, to do so. Indeed, most publishers have few incentives to publish users' speech in the face of even mild government pressure to censor it. As this Court and lower courts have recognized, "The distributor who is prevented from selling a few titles is not likely to sustain sufficient economic injury to induce him to seek judicial vindication of his rights. The publisher has the greater economic stake, because suppression of a particular book prevents him from recouping his investment in publishing it." *See Bantam Books*, 372 U.S. at 64 n.6. *See also Backpage.com v. Dart*, 807 F.3d 229, 230 (7th Cir. 2015). Speakers from marginalized communities or who hold controversial or unpopular views that do not provide intermediary publishers significant financial benefits are especially victimized by passive acquiescence to government censorship demands.

As will be discussed below, content moderation systems are fraught, and governmental manipulation of them to control public dialogue, silence disfavored voices, or blunt social movements raises classic First Amendment

concerns. Governments have outsized influence to manipulate sites' content moderation processes for their own political goals. This is especially problematic for government actors such as law enforcement whose communications may be inherently threatening or intimidating.

B. Government May Productively and Appropriately Contribute to Content Moderation Practices

Despite the concerns presented by governments' involvement in content moderation processes, not every government communication to a social media site is either improper or unwise. Users generally want to receive authoritative and accurate information that they can trust, situated within a healthy information environment. Interactions between sites and governments can be important means of fostering healthier and more reliable information environments on social media. And they can also diminish the impact of disinformation, spam, and other efforts to distort the marketplace of ideas.

The government is especially well-poised to contribute to trustworthy information environments on topics where it holds expertise. For example, voters may want to receive accurate information about the time, place, manner, and qualifications to vote; the most reliable source for that information is often their local election official. People may also want the most current information about weather emergencies, natural disasters, and other events affecting public safety; government officials often are the most authoritative sources of that information. Government officials can also inform sites about threats to accurate

and reliable information environments, like disinformation campaigns and intelligence related to foreign actors attempting to sway elections or otherwise harm national security, that can aid in reducing or interrupting such threats.[4]

Moreover, the government may have a role in publicly criticizing social media sites when it disagrees with their content moderation actions, including when it believes that the sites' policies or practices are advancing their private interests at the expense of the public interest.

Governments at all levels have the authority to advance their viewpoints and attempt to convince others to adopt them. Indeed, the federal government does this routinely, as this Court's prior cases have noted. The National Endowment for Democracy, for example, was established for the United States to encourage other countries to adopt democratic principles. *See* 22 U.S.C. § 4411(b); *Rust v. Sullivan*, 500 U.S. 173, 194 (1991). And the former Safe and Drug-Free Schools and Communities Act of 1994, 20 U.S.C. § 7114(d)(6) required schools receiving federal funds to have programs that "convey a clear and consistent message that … the illegal use of drugs [is] wrong and harmful." *See Morse v. Frederick*, 551 U.S. 393, 408 (2007).

A current example is the National Highway Transportation Authority's campaign against "Buzzed

4. AJ Vicens, *Ahead of 2024 election, Meta worries about lack of information on top-tier nation-state covert operations*, Cyberscoop (Nov. 30, 2023), https://cyberscoop.com/ahead-of-2024-election-meta-worries-about-lack-of-information-on-top-tier-nation-state-covert-operations/.

Driving," that is, driving after having consumed "[e]ven a small amount of alcohol," when one's blood alcohol content is below the legal limit of .08 and thus noncriminal in most states.[5] NHTSA is not disempowered from publicly admonishing those speakers who advocate *for* buzzed driving. NHTSA can criticize the speakers, call their information out as being false, and even encourage those who have published contrary information to depublish it; but the agency may not coerce or otherwise improperly demand censorship. A social media site that believed NHTSA's message was the better one and wanted to minimize posts on its site that encouraged "dangerous behavior" might choose to moderate the posts called out by NHTSA. A site might also consider posts that asserted that buzzed driving is safe to be in violation of their misinformation or public safety policies. A different site might disagree with the NHTSA's position and choose to amplify the contrary user posts.

II. FIRST AMENDMENT DOCTRINE REFLECTS THESE COMPETING INTERESTS AND REQUIRES COURTS TO DISTINGUISH BETWEEN APPROPRIATE AND INAPPROPRIATE GOVERNMENT COMMUNICATIONS

Courts have the difficult task of distinguishing between the government's permissible, hortatory communications—notifying, advising, persuading, convincing, mildly encouraging—and the impermissible, regulatory ones—coercing, threatening, compelling,

5. NHTSA, *Buzzed Driving is Drunk Driving*, https://www.nhtsa.gov/campaign/buzzed-driving (last visited July 26, 2023).

intimidating, and other communications designed not merely to change the publisher's mind but to cause "compliance" that "was not voluntary." *See Bantam Books*, 372 U.S. at 68.

The competing interests discussed above require a highly specific, contextual, totality of the circumstances, case-by-case analysis. Outside of blatant coercion under the threat of criminal prosecution for noncompliance, there are unfortunately few bright lines. This Court must clarify the analysis and provide sufficient guidance to allow appropriate and useful communications to resume.

A. *Bantam Books* Provides the Proper Framework For Assessing Improper Government Inducement of Private Censorship

The proper test to assess whether the government has improperly effectuated censorship originates from *Bantam Books*, in which this Court found that the First Amendment prohibited not only direct censorship demands but also "system[s] of informal censorship" aimed at speech intermediaries. 372 U.S. at 71. The Supreme Court found that "the threat of invoking legal sanctions and other means of coercion, persuasion, and intimidation" against book distributors were enough to violate the First Amendment rights of the publishers of the targeted books. *Id*. at 67.

In *Bantam Books*, a state commission engaged in such unconstitutional "informal censorship." The commission issued notices to book distributors that "certain designated books," published by plaintiffs, were "objectionable for sale," and that it was the commission's

"duty to recommend to the Attorney General prosecution of purveyors of obscenity." *Id.* at 61-62. The commission also circulated the notices to local police, who visited the distributor "to learn what action he had taken." *Id.* at 62-63. These notices were "phrased virtually as orders, reasonably understood to be such by the distributor" and, predictably, the distributors stopped selling the books. *Id.* at 64, 68.

The Court found that the publishers had a First Amendment remedy against the state commission, even though it was the distributor's action that directly harmed the publishers' sales, and the government did not actually seize any books or prosecute anyone. *Id.* at 64 & n.6. The Court found that the commission intended to censor—it had "deliberately set about to achieve the suppression of publications deemed 'objectionable'"—and that the distributor's decision to pull the books "was not voluntary." *Id.* at 67-69.

Cases following *Bantam Books* have applied a totality of the circumstances analysis and identified numerous factors relevant to determining when government improperly pressured a speech intermediary to involuntarily censor other speakers. Each factor helps courts analyze (1) whether the government "deliberately set about to achieve the suppression of publications deemed objectionable," *id.* at 67; and (2) whether a reasonable person in the shoes of the target of the communications would perceive compliance with the government's request to be truly optional or effectively required. This Court should view all factors surrounding a communication as relevant—with no single factor determinative of the outcome—and weighted in favor of protecting First Amendment rights.

In *Backpage.com v. Dart*, the Seventh Circuit followed *Bantam Books* to distinguish between "attempts to convince and attempts to coerce," the former being permissible and the latter forbidden. 807 F.3d 229, 230 (7th Cir. 2015). The case involved a sheriff's campaign to shutdown Backpage.com's adult section "by demanding that firms such as Visa and MasterCard prohibit the use of their credit cards to purchase any ads on Backpage." *Id.* The sheriff demanded in writing, on official letterhead, that the credit card companies "cease and desist" allowing payments for Backpage ads, citing the federal money-laundering statute. *Id.* at 231-32.

In deciding that the sheriff's letter to the credit companies was unconstitutional under *Bantam Books*, the court considered several factors.

First, the letter constituted an "implied threat" because it was written in the sheriff's official capacity, "invoke[ed] the legal obligations of financial institutions to cooperate with law enforcement," and required further ongoing contact from the companies following the request. *Id.* at 236. The letter was not simply an effort to educate the recipients "about the nature and possible consequences of advertising for sex; he told them to desist or else." *Id.* at 237. Because of the threatening language, the letter was not "a permissible attempt at mere persuasion." *Id.* at 238.

Second, as in *Bantam Books*, Sheriff Dart clearly *intended* to coerce Visa and Mastercard to cooperate. A strategic memo recommended appealing to the intermediaries' interest in avoiding liability; the sheriff took credit in a press release for "compelling" the companies' actions with his "demand"; and his office sent

urgent communications to the companies following up on his letter which "imposed another layer of coercion due to [their] strong suggestion that the companies could not simply ignore [the sheriff]." *Id.* at 232, 237.

Finally, though it was not a necessary finding,[6] the court considered that the letter achieved its censorial goals. Visa and Mastercard had each received similar complaints from private citizens in the past, but severed ties with Backpage only two days after receiving the sheriff's letter. *Id.* at 232-33. And even though Visa denied that it felt threatened, the court found an "obvious" causality between the sheriff's letter and the credit card companies' decisions to comply. *Id.* at 233.

Looking to other courts, the Ninth and Second Circuits have identified four relevant factors, with no single one being dispositive: (1) word choice and tone, and "the tenor of the overall interaction"; (2) the existence of regulatory authority to punish noncompliance; (3) whether the speech was perceived as a threat; and, perhaps most importantly, (4) whether the speech refers to adverse consequences. *Kennedy v. Warren*, 66 F.4th 1199, 1207-09 (9th Cir. 2023).[7]

---

6. Indeed, the Seventh Circuit emphasized that "such a threat is actionable and thus can be enjoined even if it turns out to be empty—the victim ignores it, and the threatener folds his tent." *Id.* at 231.

7. The Second Circuit's decision, which derived these four factors from that court's previous cases, is currently being reviewed by this Court. *See Nat'l Rifle Ass'n of Am. v. Vullo*, 49 F.4th 700, 715 (2d Cir. 2022), *cert. granted*, ___S.Ct. ___, 2023 WL 7266997 (Mem) (Nov. 3, 2023).

Other courts have found numerous additional factors relevant to answering the questions regarding speaker intent and recipient perception.

Courts have considered whether the communication contained a threat of adverse consequences, and if so, whether they were legal or economic. If legal, the court may inquire whether the government threatened criminal prosecution or merely a routine administrative action.[8] For economic threats, courts may ask whether an intermediary in some way relied on the government's favorable opinion for preference in securing future contracts, advantageous legislation, or some other public benefit.[9]

Courts tend to find government communications permissibly hortatory rather than regulatory when the official affirmatively and genuinely disclaims any authority or intent to sanction.[10] Courts similarly tend to find communications permissible when they include

8. *See R.C. Maxwell Co. v. Borough of New Hope*, 735 F.2d 85, 88 (3d Cir. 1984) ; *Penthouse Int'l, Ltd. v. Meese*, 939 F.2d 1011, 1015 (D.C. Cir. 1991).

9. *See Okwedy v. Molinari*, 333 F.3d 339, 343 (2d Cir. 2003); *Rattner v. Netburn*, 930 F.2d 204, 210 (2d Cir. 1991); *R.C. Maxwell*, 735 F.2d at 87.

10. *See R.C. Maxwell*, 735 F.2d at 86 n.2 ("We are writing to solicit your personal assistance in order to alleviate an ongoing situation in our community by a professional agreement rather than legal procedures."); *VDARE Found. v. City of Colorado Springs*, 11 F.4th 1151, 1157 (10th Cir. 2021), *cert. denied*, 142 S. Ct. 1208 (2022) (finding no coercion where the mayor issued public statement against hate speech and encouraged businesses to be attentive to the "types of events they accept" but acknowledged his lack of authority to restrict freedom of speech).

a reasoned argument for the censorship rather than a threat of penalty.[11]

Courts are also more likely to find communications impermissibly regulatory when the government identifies specific objectionable speech, and find no wrongdoing when the government generally disapproves of a broad category of speech or an intermediary's general practices.[12]

Also pertinent as to both intent and perception is whether the government took further action after the initial request.[13] Immediate and serious follow-up communications "continually reinforce[]" the request and thus may cause the intermediary to reasonably interpret it as mandatory. *VDARE*, 11 F.4th at 1167.

Government requests through a channel of communication created by the intermediary, particularly where the intermediary welcomed or solicited the government's input or expertise, are more likely to be

---

11. *Hammerhead Enterprises, Inc. v. Brezenoff*, 707 F.2d 33, 36 n.2 (2d Cir. 1983), *cert. denied,* 464 U.S. 892 (1983) (finding no coercion where letter stated "Your cooperation in keeping this game off the shelves of your stores would be a genuine public service"); *Kennedy*, 66 F.4th at 1204 (finding no coercion in absence of threat).

12. *See VDARE*, 11 F.4th at 1165 (citing *R.C. Maxwell*, 735 F.2d at 86) (finding no coercion in city statement that generally "encourage[d] local businesses to be attentive to the types of events they accept and the groups that they invite to our great city" that did not specifically name the plaintiff).

13. *Kennedy*, 66 F.4th at 1209 ("An interaction will tend to be more threatening if the official refuses to take 'no' for an answer and pesters the recipient until it succumbs.").

permissible.[14] Also relevant is whether the communication was part of a public announcement, which because of their transparency lean hortatory, or through a private channel, which lean regulatory. *Compare Kennedy*, 66 F.4th 1199 (finding Senator's open letter hortatory) *with Rattner*, 930 F.2d at 209 (finding official's private letters potentially regulatory).

Lastly, courts are likely to view interactions with power imbalances between the government actor and the targeted intermediary—such as where law enforcement is involved[15]—as more conducive to bullying or intimidation.[16]

---

14. *See O'Handley v. Weber*, 62 F.4th 1145, 1163 (9th Cir. 2023) (finding no coercion where "OEC communicated with Twitter through the Partner Support Portal, which Twitter voluntarily created because it valued outside actors' input").

15. "[T]he emerging tactic of law enforcement officials in targeting ISPs with 'requests' that they take down websites that officials find problematic raises, in modern form, the threats to free expression implicit in any mechanism of prior restraint." Seth F. Kreimer, *Censorship by Proxy: The First Amendment, Internet Intermediaries, and the Problem of the Weakest Link*, 155 U. Pa. L. Rev. 11, 76–77 (2006); *see also Kennedy*, 66 F.4th at 1210 ("A similar letter might be inherently coercive if sent by a prosecutor with the power to bring charges against the recipient, or if sent by some other law enforcement officer[.]"); *Rattner*, 930 F.2d at 209 ("unannounced visits by police personnel" are relevant to consideration of an implied threat).

16. *Speech First, Inc. v. Cartwright*, 32 F.4th 1110, 1123 (11th Cir. 2022) ("[T]he students targeted here are—for the most part—teenagers and young adults who, it stands to reason, are more likely to be cowed by subtle coercion than the relatively sophisticated business owners in those cases."); *see also* Derek E. Bambauer, *Against Jawboning*, 100 Minn. L. Rev. 51, 103–106 (2015).

B. The State Action Coercion Doctrine Should Be Read As Coextensive With the First Amendment Analysis Under *Bantam Books*.

The distinction between permissible and impermissible government communications is made even fuzzier by the uncertain interaction of the *Bantam Books* analysis with the coercion analysis from the state action doctrine, which some courts employ instead of or in addition to *Bantam Books*.

To bring much needed clarity, this Court should hold that *Bantam Books* and state action coercion are simply the same analysis, with the same relevant factors and limitations applying to each.

The state action coercion analysis, from *Blum v. Yaretsky*, provides that the government is liable for a private entity's censorship decision when the government coerces or provides "such significant encouragement" to that private entity that its decision must be deemed the government's. 457 U.S. 991, 1004 (1982).

Blum thus requires courts to distinguish between permissible "encouragement" and impermissible "such significant encouragement." Unfortunately, neither *Blum* nor its progeny provide much explanation or guidance about how to identify these impermissible communications, thus exacerbating the fuzziness of the analysis.

Blum's coercion test is best read as being essentially the same as the *Bantam Books* test, with *Blum*'s "such significant encouragement" the equivalent of *Bantam*

Books's implied censorship. Each seeks to answer the same question: is the decision of the private intermediary to censor another's speech legally attributable to the state?

This Court should clarify that the numerous factors laid out above apply to, and cabin, both analyses. Indeed, lower courts have recognized that the *Bantam Books* and *Blum* analyses share a common central inquiry in suits alleging state coercion of private censorship and relied upon both analyses. In *VDARE*, the Tenth Circuit used *Blum* as a general framework for coercion, and *Bantam Books* and its progeny as a specific application of coercion in the First Amendment context. *See* 11 F.4th at 1159–1168. And the Ninth Circuit's *Bantam Books* analysis in *Kennedy* relied upon state action cases—namely, *Carlin Communications, Inc. v. Mountain States Tel. & Tel. Co.*, 827 F.2d 1291 (9th Cir. 1987), and *VDARE*. *See* 66 F.4th at 1207–12.

The two tests commonly reach the same result. For example, similar to the Seventh Circuit's *Bantam Books*-based decision in *Dart,* the Ninth Circuit in *Carlin* found coercion under the state action doctrine on almost identical facts: a local government attorney sent a letter to a speech intermediary threatening to prosecute it for transmitting messages with sexual content, citing a state statute prohibiting display of sexually explicit material to minors. *Carlin*, 827 F.2d at 1923, 1925.

C. Close Cases Should Be Resolved Against the Government Since the Government Is Best Able to Model its Communications to Avoid Ambiguity

In the ample grey area that is "informal censorship"/"such significant encouragement," the government should ultimately bear the onus of making sure that it is clearly not trying to replace the company's editorial decision-making with its own. Close cases should thus be resolved in favor of the unjustly censored party. Such a rule will incentivize the government to take care that its communications are clearly perceived as being merely informative or persuasive but not threatening, intimidating, or coercive. When government speaks to speech intermediaries about the speech they publish, it must do so with great care and with a keen awareness of its inherent coercive power and potential for intimidation.

Of all the parties participating in such communications, government speakers are best placed to bear this responsibility. They are also the only parties to the communication bound by the First Amendment. The government may easily start with training government speakers to include affirmative disclaimers—for example, "while we believe this is the right thing to do, the ultimate decision remains yours and you will not be penalized for making your own decision"—in pertinent communications.

The federal government should also provide more transparency into its interactions with social media sites. When possible, the government should make the communication itself public. In others, the government should disclose information about its communications with

sites, including, for example, subject and timing, without compromising privacy. These would give the public an opportunity to evaluate the government's involvement in the content moderation process and challenge the government's actions if desired.

Moreover, there is a diminished risk of chilling effect on government speakers. Government speech tends to be quite hearty and politically motivated. And the fear of close-case or vexatious litigation tends not to discourage government speakers.

III. THIS COURT'S ANALYSIS MUST BE INFORMED BY A FULL UNDERSTANDING OF CONTENT MODERATION PROCESSES

Because context is an important part of the First Amendment analysis, this Court must understand how the often quixotic content moderation processes of social media sites operate in order to determine whether government has exercised improper influence in any particular case. Of particular relevance here, the notices, complaints, and take-down requests from the government are among the millions of complaints, reports, and "flags" that social media companies routinely receive and sometimes act on from a variety of stakeholders. That a site seeks out, receives, or considers opinions on certain user speech from the government is therefore not in and of itself exceptional, and does not signal that the government was the only stakeholder engaged with the site on that issue. A court might thus consider which of the site's departments the government communicated with, whether it was through an established process like a trusted flaggers program, and whether other non-governmental stakeholders were involved in the advocacy.

A. Content Moderation Is an Historic and Widely Employed Practice With Varying Standards and Practices

Because this Court has received ample explanations of the content moderation process from the briefs filed in the pending cases, *Netchoice v. Paxton*, No. 22-255 and *Moody v. Netchoice*, No. 22-277, this brief provides only a high-level overview.

Social media sites, at least from their point of mass adoption, have rarely published all legal speech submitted to their sites. Instead, they engage in content moderation: the use of policies, systems, and tools to decide what user-generated content or accounts to publish, remove, amplify, or manage.[17]

Sites practice content moderation in phases: they define permissible and impermissible content via content and usage policies; detect content that may violate their policies or the law; evaluate that content to determine whether it in fact violates their policies or the law; take an enforcement action against violative content; and often allow users to seek review of content moderation decisions that they believe are erroneous.[18] In each phase, sites make editorial judgments about what content they wish to allow or forbid on their services, or how to display or arrange it.

---

17. *See* Hannah Bloch-Wehba, *Automation in Moderation*, 53 Cornell Int'l L.J. 41, 42–43, 48 (2020).

18. Seny Kamara et al., *Outside Looking In: Approaches to Content Moderation in End-to-End Encrypted Systems*, Ctr. for Democracy & Tech., 9–11 (2021), https://cdt.org/wp-content/uploads/2021/08/CDT-Outside-Looking-In-Approaches-to-Content-Moderation-in-End-to-End-Encrypted-Systems-updated-20220113.pdf.

Content moderation differs from site to site.[19] Some sites detect potentially violating content only after it is posted; others screen some or all content *ex ante*.[20] Sites make different judgment calls about whether particular content violates their content policies, even if those policies are similar.[21] They use different methods to enforce their content policies, not only removing violative posts, but also potentially changing the manner and place in which posts are displayed, preventing monetization, or adding the site's own affirmative speech.[22] Some sites allow users to appeal content moderation decisions, while others do not.[23]

Many users prefer moderated sites. Users may want to find or create affinity and niche communities dedicated to certain subject matters or viewpoints and exclude others.

---

19. *Compare Community Guidelines,* Instagram, https://help.instagram.com/477434105621119 (last visited December 15, 2023) (prohibiting nudity), *with Sensitive Media Policy*, Twitter (March 2023), https://help.twitter.com/en/rules-and-policies/media-policy (permitting "consensually produced adult nudity").

20. Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 Harv. L. Rev. 1598, 1635 (2017).

21. *See, e.g.,* Hannah Denham, *Another Fake Video of Pelosi Goes Viral on Facebook*, Wash. Post (Aug. 3, 2020), https://www.washingtonpost.com/technology/2020/08/03/nancy-pelosi-fake-video-facebook/ (reporting that TikTok, Twitter and YouTube removed a doctored video of Rep. Nancy Pelosi, while Facebook allowed it to remain with a label).

22. *See* Eric Goldman, *Content Moderation Remedies*, 28 Mich. Tech. L. Rev. 1, 23–39 (2021) (describing various enforcement options).

23. Klonick, *supra* n.20, at 1648.

They may prefer environments that shield them from certain kinds of legal speech, including misinformation, hateful rhetoric and harassment, or simply speech that is off-topic or irrelevant.[24] And all users want services to filter out junk content or "spam."

Users may want a service that only has highly trustworthy information and actively attempts to filter out misinformation. For example, one of the sites specified in the district court's order, D.Ct Judgment, 3 n.2, Pinterest, a site designed to visually inspire creative projects, has "community guidelines" that "outline what we do and don't allow on Pinterest."[25] Among the prohibited categories is "Misinformation."[26] Another site specified in the district court's order, YouTube, *see* D.Ct Judgment, 3 n.2, prohibits "misinformation" with serious risk of egregious harm.[27] Moreover, its policy prohibiting promotion or glorification of Nazi ideology specifically prohibits misinformation in the form of Holocaust denial.[28]

---

24. *See, e.g.*, *Reducing Hate And Disinformation Online*, Change the Terms, https://www.changetheterms.org (last visited December 15, 2023) (campaign demanding improved content moderation against hate speech and disinformation).

25. *Community Guidelines*, Pinterest, https://policy.pinterest.com/en/community-guidelines (last visited December 15, 2023).

26. *Id.*

27. *Misinformation Policies*, YouTube Help, https://support.google.com/youtube/answer/10834785 (last visited December 15, 2023).

28. The YouTube Team, *Our Ongoing Work to Tackle Hate*, YouTube (June 5, 2019), https://blog.youtube/news-and-events/our-ongoing-work-to-tackle-hate/.

Content moderation also helps users avoid misinformation in the form of spam and scams. For instance, employment websites that allow employers to post job openings use spam and scam policies to combat, among other things, a growing trend of scammers using employment websites to steal applicants' identities in order to commit unemployment benefit fraud.[29] LinkedIn, for example, one of the sites specified in the district court's order, D.Ct Judgment, 3 n.2, removes "phishing links, malware, known or suspected scam content, and fraudulent content and permanently restrict the accounts of known fraudsters or scammers," pursuant to its "Professional community policies."[30] And LinkedIn encourages users who see scam postings to report them.[31]

But misinformation is far from the only content category most sites seek to exclude. Many sites use content moderation to create environments they believe are more user-friendly, including being focused on specific interests, and prohibit content that the sites deem unsuitable for their purposes. For example, Peanut, a social media site

29. *See* Cezary Podkul, *Scammers Are Using Fake Job Ads to Steal People's Identities*, ProPublica (Oct. 26, 2021), https://www.propublica.org/article/scammers-are-using-fake-job-ads-to-steal-peoples-identities.

30. *LinkedIn Professional Community Policies*, LinkedIn, https://www.linkedin.com/legal/professional-community-policies (last visited December 15, 2023); *Scams and Fraud Content*, LinkedIn, https://www.linkedin.com/help/linkedin/answer/a1338803 (last visited December 15, 2023).

31. Muhammed Imran Shafique, *How to Spot and Avoid LinkedIn Job Scams*, LinkedIn (Nov. 26, 2022), https://www.linkedin.com/pulse/how-spot-avoid-linkedin-job-scams-imran-shafique.

aiming to be a "safe, inclusive space for women" navigating fertility, pregnancy, and motherhood similarly prohibits its millions of users from posting any content that attacks, threatens or "otherwise dehumanizes an individual or group" based on race, religion, age, socioeconomic status, or disability, among other categories.[32] In addition to "misinformation," "bullying," and "nudity, pornography, sexually explicit content or sexual solicitation" are also barred.[33] Peanut additionally prohibits users from "raising money on behalf of other individuals," requesting "financial aid," or promoting their own gift "wish lists."[34]

Content moderation is a difficult and often fraught process that even the largest and best-resourced social media companies struggle with, often to the frustration of users. Even when using a set of precise rules or carefully articulated "community standards," moderated sites often struggle to draw workable lines between permitted and forbidden speech. Every online forum for user speech, not just the dominant social media sites, struggles with this problem. And every social media user has likely experienced it, either as a creator or reader.

Content moderation controversies and government involvement in them are not a new phenomenon. In 2007, YouTube, only two years old at the time, shut down the account of Egyptian human rights activist Wael Abbas after receiving multiple reports that the account featured

---

32. *Community Guidelines*, Peanut, https://www.peanut-app.io/community-guidelines (last visited Nov. 17, 2023).

33. *Id.*

34. *Id.*

graphic videos of police brutality and torture.[35] YouTube's community standards at the time stated that "[g]raphic, or gratuitous violence is not allowed."[36] Abbas's account was restored only after the U.S. State Department communicated with YouTube's new owner, Google.[37]

B. Social Media Sites Commonly Consult With and Receive Advice and Feedback From Numerous External Sources, Including Governments

For users to even hope for fairness and consistency in content moderation decisions, social media sites need to draw on outside resources and expertise. This practice, which includes the use of trusted flagger programs, trust and safety councils, external stakeholder engagement teams, as well as as-needed consultations with individual and organizational experts, is widespread and often referred to as "networked governance."[38]

Governments are thus among many entities and individuals that notify, contact, urge, or encourage the various social media companies to moderate user posts. And governments are also among the entities and individuals to which the sites will themselves reach out when they are seeking expertise.

---

35. Kevin Anderson, *YouTube Suspends Egyptian Blog Activist's Account*, The Guardian (Nov. 28, 2007), https://www.theguardian.com/news/blog/2007/nov/28/youtubesuspendsegyptianblog.

36. *Id.*

37. Jillian C. York, *Silicon Values: The Future of Free Speech Under Surveillance Capitalism* 25-27 (Verso 2021).

38. Robyn Caplan, *Networked Governance*, 24 Yale J.L. & Tech. 541, 542 (2022).

Meta, for example, reports that it uses stakeholder engagement to both develop and implement its community standards. With respect to misinformation, Meta states that "we engage extensively with experts and civil society stakeholders on topics such as state media, harmful health misinformation, and misinformation that may contribute to a risk of offline harm," and, "Our team regularly speaks with academics and NGOs to provide visibility into how we develop and apply our policies in these areas."[39]

TikTok has "Safety Partners," with which it "share[s] best practices, create[s] programs, and exchange[s] ideas on safety-related topics." This includes its "Content Advisory Council" whose members have expertise in "child safety, hate speech, misinformation and bullying" and with whom TikTok works "to gain unvarnished views on and advice around our policies and practices as we continually work to improve them."[40]

Sites seek and receive feedback in a number of ways.

Most sites allow users to report or "flag" content they believe violates the sites' rules or standards.[41] Indeed, such flags may account for a large amount of content moderation decisions. In the third quarter of 2023, YouTube removed

---

39. Meta, *How Stakeholder Engagement Helps Us Develop the Facebook Community Standards*, Meta (Jan. 18, 2023), https://transparency.fb.com/en-gb/policies/improving/stakeholders-help-us-develop-community-standards/.

40. *Safety Partners*, TikTok, https://www.tiktok.com/safety/en-us/safety-partners/ (last visited December 15, 2023).

41. *See generally* Kate Crawford & Tarleton Gillespie, *What Is a Flag For? Social Media Reporting Tools and the Vocabulary of Complaint*, 18 New Media & Soc'y 410 (2014).

8.1 million videos, 322,000 of which users flagged.[42] During that same period, Facebook users reported 12.2% of the 8.3 million posts Facebook actioned for bullying and harassment was reported by users and 5.2% of the 9.6 million posts actioned for hate speech.[43] And in the second half of 2021, users reported over 11 million accounts to Twitter as having violated at least one of its rules.[44]

Sites also commonly seek and/or receive input from civil society groups, activists, and other stakeholders who are not necessarily their users. This input may urge sites to take content down *or* to put it back up. Public interest organizations have sent public and private letters to urge social media companies to address election-related misinformation; hate speech against protected communities; content promoting anorexia and disordered eating; and more.[45]

---

42. *YouTube Community Guidelines Enforcement*, Google, https://transparencyreport.google.com/youtube-policy/removals (last visited December 15, 2023).

43. *Community Standards Enforcement Report*, Meta, https://transparency.fb.com/reports/community-standards-enforcement/ (last visited December 15, 2023).

44. *Rules Enforcement*, Twitter (July 28, 2022), https://transparency.twitter.com/en/reports/rules-enforcement.html#2021-jul-dec.

45. For example, *amicus* EFF joined efforts pressuring Facebook to end its ban on pictures of female nipples. Kari Paul, *Naked Protesters Condemn Nipple Censorship at Facebook Headquarters*, The Guardian (June 3, 2019), https://www.theguardian.com/technology/2019/jun/03/facebook-nude-nipple-protest-wethenipple.

Sites may also reach out to government agencies or officials when they perceive them as being authoritative sources of information. For example, the district court found it significant that Facebook solicited the opinion of the Centers for Disease Control regarding the accuracy of posts promoting ivermectin in treating COVID. *See Missouri v. Biden*, No. 3:22-CV-01213, 2023 WL 4335270, at *20–22 (W.D. La. July 4, 2023).

Many sites have "trusted flagger" programs in which certain groups and individuals enjoy "some degree of priority in the processing of notices, as well as access to special interfaces or points of contact to submit their flags." Naomi Appelman & Paddy Leerssen, *On 'Trusted' Flaggers*, 24 Yale J.L. & Tech. 452, 453 (2022).

YouTube's "Priority Flagger" program prioritizes complaints from government agencies and NGOs, that are "particularly effective" at notifying YouTube of violative content.[46] From July to September of 2023, the site removed 40,492 videos due to these complaints.[47]

YouTube includes government agencies among those who may be "priority flaggers";[48] in the third quarter of this year, eleven videos were removed as a result of government flags, as compared to over 400,000 takedowns originating from non-governmental external flags.[49]

---

46. *About the YouTube Priority Flagger Program*, YouTube, https://support.google.com/youtube/answer/7554338 (last visited December 15, 2023).

47. Google, *supra* n.42.

48. YouTube, *supra* n.46.

49. Google, *supra* n.42.

Of course, governmental participation as priority "flaggers" raises First Amendment issues not present with non-governmental inputs, since the designation may create undue pressure over sites' content moderation decisions. As just one recent example of the hazards of establishing flagging pathways between governments and sites, the Meta Oversight Board reviewed the decision by Meta to remove a musical track from Instagram, after Instagram received a request from London Metropolitan Police to review all content containing the track because the song, the police opined, contained a veiled threat of gang violence and could lead to further violence. [50] Ultimately, 165 pieces of content containing the track were removed.[51] Through freedom of information laws, the Oversight Board learned that the police had made 286 requests to various social media companies to remove tracks of the same genre, drill music, resulting in 255 removals across various sites.[52] The Oversight Board overturned the decision, and also recommended that Meta increase transparency around government take-down requests and regularly review data on content moderation decisions prompted by government requests for systemic bias.[53]

---

50. Meta Oversight Board, *UK Drill Music* (Jan. 2023), https://oversightboard.com/decision/IG-PT5WRTLW/.

51. *Id.*

52. *Id.*

53. *Id.*

IV. EXAMINING SOME OF PETITIONERS' INVOLVEMENT IN CONTENT MODERATION

As with all First Amendment issues, this Court should independently review the record before deeming any particular communication to have crossed the constitutional line. *See Bose Corp.*, 466 U.S. at 499-500. Independent review is especially important in this case because it has been reported that the lower courts' factual characterizations were not supported by the record.[54] Moreover, because the lower court and the Fifth Circuit failed to provide proper guidance to the government regarding its interactions with the sites, the government may be refraining from permissible conduct, interrupting information flows that have been helpful to sites and ultimately to users.

Amici offer a few examples of how a correctly applied constitutional test might play out. As these examples show, some communications at issue appear to be permissible steps to share information or to persuade, rather than coerce, the recipient. And there are instances where the government likely crossed the line into impermissible coercion.

For example, Rob Flaherty's communications to Facebook regarding specific Tucker Carlson and Tomi

54. *See, e.g.*, Mike Masnick, *5th Circuit Cleans Up District Court's Silly Jawboning Ruling About the Biden Admin, Trims It Down To More Accurately Reflect The 1st Amendment*, Techdirt (Sep 11, 2023, 09:31 AM), https://www.techdirt.com/2023/09/11/5th-circuit-cleans-up-district-courts-silly-jawboning-ruling-about-the-biden-admin-trims-it-down-to-more-accurately-reflect-the-1st-amendment/.

Lahren posts expressing COVID-19 vaccine hesitancy, *see* D.Ct. Doc. 174, Attachment 1, at 22, presents at least a close case that should likely be resolved against the government.

Flaherty served as Deputy Assistant to the President and Director of Digital Strategy, and an objective observer would reasonably think Flaherty had a role in setting, implementing, or enforcing administration policy based on his position and communications. One email Flaherty sent to Facebook—containing government findings concerning COVID-19 vaccine misinformation on the site—includes his comment "Don't read this as White House endorsement of these suggestions (or, also, as the upper bound of what *our* thoughts on this might be). But – spirit of transparency – this is circulating around the building and informing thinking." D.Ct. Doc. 214, Attachment 14, at 1 (emphasis added). Another email Flaherty sent to Facebook states "We are concerned that your service is one of the top drivers of vaccine hesitancy- period," and "we want to know that you're trying, we want to know how we can help, and we want to know that you're not playing a shell game with us when we ask you what is going on." D.Ct. Doc. 174, Attachment 1, at 11.

At the time of the Carlson and Lahren posts, Facebook's policy was to reduce the distribution of content that contributed to "unfounded hesitancy" regarding COVID-19 vaccines so that fewer people would see such content. D.Ct. Doc. 174, Attachment 1, at 5. Facebook communicated this policy in a February 9 reply email to Flaherty. *See id.* at 4–5, 7.

On April 14, a Flaherty email to Facebook stated that "the top post about vaccines today is tucker [sic] Carlson saying they don't work," and "[y]esterday was Tomi Lehren [sic] saying she won't take one." *Id.* at 22. Flaherty's email continued: "This is exactly why I want to know what 'Reduction' actually looks like – if 'reduction' means 'pumping our most vaccine hesitant audience with Tucker Carlson saying it does not work'… then…I'm not sure it's reduction!" *Id.* A second White House official also emailed Facebook regarding the same Carlson post, stating "Number one on Facebook. Sigh." D.Ct. Doc. 174, Attachment 1, at 35.

Later that day, Facebook included Flaherty in its response to the second White House official. Facebook's response explained that, while "the Tucker Carlson video does not qualify for removal under [Facebook's] policies," the site was labelling, demoting, and not recommending the video. *Id.* at 34.

Flaherty replied that evening by email asking how the content did not violate Facebook's policies, Facebook's "rule for removal vs demoting," and what Facebook meant when it discussed reducing and demoting. *Id.* at 33. The message concluded: "Not for nothing but last time we did this dance, it ended in an insurrection." *Id.* at 34. Flaherty would email Facebook two more times regarding the same Carlson and Lahren posts over the next two days. *Id.* at 33.

It may be appropriate for the government to request that a site explain how its moderation policies apply to specific content. It may be appropriate for the government to express its disagreement, publicly or privately, with how a site has applied its moderation policies to specific content.

And it may also be appropriate for the government to offer a reasoned argument to a site based on its moderation policies in an attempt to persuade the site to come to a different enforcement conclusion.

But the Flaherty communications, exhibiting through their word choice, tone, and context a government actor's intent to pressure Facebook to action specific content, contrary to Facebook's own determination, are at least in the grey area between a permissible attempt to persuade and impermissible censorship. Their coercive nature is underscored by the context in which they occurred—the larger nonpublic, dialogue between the White House and Facebook concerning the same specific posts. The totality of the circumstances, including the repeated inquiries about and frustration over why Facebook did not action specific content in the way the government preferred appear more like brow-beating than a reasoned attempt at persuasion.

In contrast, the Centers for Disease Control and Prevention appear to have permissibly responded to Facebook requests that the agency provide its opinions regarding the accuracy of certain claims about COVID-19 vaccines' side-effects. *See Missouri*, 2023 WL 4335270, at *20–22. Certainly, the First Amendment does not prevent the CDC from providing its opinion when asked. The CDC may respond to such inquiries with its positions—even knowing that Facebook would rely upon them in moderating its site, *id.*—where the record gives no indication that these CDC responses contained any subtle or not-so-subtle threats of adverse consequences. Such communications appear to have been "reasoned arguments" that courts have found persuasive rather

than coercive. Moreover, the CDC lacks any enforcement authority over Meta.

Lastly, amici address the allegations against the FBI since law enforcement communications raise special concerns because of their inherently coercive nature. *See, e.g., Bantam Books*, 372 U.S. at 68–69; *Dart*, 807 F.3d at 236–37; *Carlin*, 827 F.2d at 1296.

Here, the FBI's communications with sites appear to be permissibly informative. The FBI sought to understand the sites' processes and provide them with pertinent information but not dictate that their editorial judgments be replaced with those of the government. The record indicates that in many of the FBI's communications with sites, the agency communicated only that it anticipated disinformation operations, *see* D.Ct. Docket 204, Attachment 1, at 154, 155–57, 172–78; D.Ct Docket 204, Attachment 5, at 2–3; D.Ct Docket 204, Attachment 6, at 56, or posed "hypotheticals" regarding how terms of service *would* apply to disinformation or misinformation. *See* D.Ct. Docket 204, Attachment 1, at 247–251.

The district court took issue with the FBI's following up with sites after flagging suspected "inauthentic" accounts,[55] *see Missouri*, 2023 WL 4335270, at *30–31, but the FBI's follow-ups appear limited to inquiring whether sites found the information the FBI provided useful. *See* D.Ct. 204, Attachment 1, at 102–03. The record does not suggest that the FBI's follow-ups challenged the adequacy of sites' responses to information the FBI provided. *See id.*

55. The definition of "inauthentic" accounts given in the record are accounts in which "the user is pretending to be someone they are not." D.Ct. Docket 204, Attachment 1, at 92.

Similarly, CISA's "switchboarding,"—forwarding to sites information that state and local election officials have identified as disinformation aimed at their jurisdictions for sites to moderate pursuant to their policies, *see Missouri*, 2023 WL 4335270, at *32—also appears permissible, despite the Fifth Circuit's finding that it supported the FBI's coercion, *Missouri v. Biden*, 83 F.4th 350, 391 (5th Cir. 2023). "Switchboarding" appears to be an example of the government using its unique resources, capabilities, and expertise to inform the sites' moderation practices.

These permissible practice serve a useful purpose for sites seeking to create healthier and more reliable information environments on social media, efforts which the lower courts' orders have inhibited. Meta recently reported that the U.S. government ceased sharing intelligence related to foreign election interference following the district court's injunction in this case, and expressed concern that, without this information, Meta will lack awareness of the bigger threat picture and have a harder time enforcing its own rules.[56]

CONCLUSION

This Court should clarify the boundaries between permissible and impermissible government communications with speech intermediaries, with *Bantam Books* as controlling doctrine and identifying relevant factors. This Court should independently examine the

56. Rebecca Kern and Maggie Miller, *Meta says US stopped sharing foreign threat intel on elections*, Politico (Nov. 30, 2023), https://subscriber.politicopro.com/article/2023/11/meta-says-us-stopped-sharing-foreign-threat-intel-on-elections-00129181.

record and carefully distinguish the proper from improper contacts alleged in this case, with an understanding of the context in which social media companies ultimately decide whether and how to take action on any of the millions users posts they publish every day.

Dated: December 21, 2023

Respectfully submitted,

David Greene
Counsel of Record
Electronic Frontier Foundation
815 Eddy Street
San Francisco, CA 94109
(415) 436-9333
davidg@eff.org

Samir Jain
Kate Ruane
Center for Democracy & Technology
1401 K Street NW,
Suite 200
Washington, DC 20005

Attorneys for Amici Curiae