

TAILORING ELECTION REGULATION: THE PLATFORM IS THE FRAME

Julie E. Cohen *

CITE AS: 4 GEO. L. TECH. REV. 641 (2020)

TABLE OF CONTENTS

| | |
|--|-----|
| I. INTRODUCTION | 641 |
| II. SQUARE PEGS, ROUND HOLES, AND FALSE IMPERATIVES: FRAMES FROM THE FIRST AMENDMENT’S PAST..... | 642 |
| A. From Distribution Bottlenecks to Microtargeting at Scale..... | 643 |
| B. From Autonomy to Automaticity..... | 649 |
| C. From “Neutral Tools” to Amplified Flows..... | 653 |
| III. FROM INFORMATION MARKETPLACES TO INFORMATION PLATFORMS: FRAMING THE FIRST AMENDMENT’S FUTURE | 656 |
| IV. CONCLUSION..... | 662 |

I. INTRODUCTION

According to conventional wisdom, legislative efforts to limit platform-based electoral manipulation—including especially laws that go beyond simply mandating additional disclosure about advertising expenditures—are most likely doomed to swift judicial invalidation for two reasons. First, although one might wonder whether the data-driven, algorithmic activities that enable and invite such manipulation ought to count as protected speech at all, the Court’s emerging jurisprudence about the baseline coverage of constitutional protection for speech seems poised to sweep many such information processing activities within the First Amendment’s ambit.¹ Second, assuming First Amendment coverage, the level

* Mark Claster Mamolen Professor of Law and Technology, Georgetown Law. Thanks to Chris Conrad for outstanding research assistance and to Jack Balkin, Ryan Calo, Erin Carroll, Danielle Citron, Jennifer Daskal, Laura Donohue, Alexandra Givens, Ellen Goodman, Margot Kaminski, Kyle Langvardt, Amanda Levendowski, Paul Ohm, David Pozen, Neil Richards, Jessica Silbey, Gigi Sohn, and Rebecca Tushnet for their helpful comments.

¹ See *Sorrell v. IMS Health Inc.*, 564 U.S. 552, 579–80 (2011); *Ashutosh Bhagwat, Sorrell v.*

of scrutiny likely to be triggered by regulation of such activities will be strict. In this Essay, I bracket questions about baseline coverage and focus on the prediction of inevitable fatality.

Legislation aimed at electoral manipulation rightly confronts serious concerns about censorship and chilling effects, but the ways that both legislators and courts approach such legislation will also be powerfully influenced by framing choices that inform assessment of whether challenged legislation is responsive to claimed harms and appropriately tailored to the interests it assertedly serves. In Part II of this Essay, I identify three frames conventionally employed in evaluating the design of speech regulation—the distribution bottleneck, the rational listener, and the intentional facilitator—and explain why each is ill-suited to the platform-based information environment, which presents different incentives and failure modes. In their place, I offer the platform itself as a new frame. Part III defines the frame more precisely, identifies the harms and interests it brings into focus, and offers some preliminary thoughts on the kinds of legislation it might permit.

II. SQUARE PEGS, ROUND HOLES, AND FALSE IMPERATIVES: FRAMES FROM THE FIRST AMENDMENT’S PAST

In any society that uses language to communicate complex ideas, frames and framing effects are inevitable. But ill-fitting frames can engender destructive feedback loops, and that is especially true where the ongoing conversation about governance within constitutionally permissible bounds is concerned. In the course of that conversation, legislators draft to the specifications of the frames they predict courts will employ without considering whether their handiwork will address the problems they want to solve. Courts then reject such efforts when the posited relationship between proposed remedies and asserted harms does not seem to make sense. Meanwhile, legislation drafted in ways more likely to be effective often dies in committee, and predictions about what courts might do with such legislation remain untested.

The three frames described below play different roles in the contemporary First Amendment landscape, but each encourages legal actors to evaluate claims about asserted dysfunctions in (real) speech environments in ways informed by certain baseline assumptions about how such

IMS Health: *Details, Detailing, and the Death of Privacy*, 36 VT. L. REV. 855, 859–61 (2012); but see Neil Richards, *Why Data Privacy Law Is (Mostly) Constitutional*, 56 WM. & MARY L. REV. 1501 (2015). I disagree that the First Amendment does or should apply to information processing activities regardless of their nature and context, but that is a subject for a different occasion. See generally Frederick Schauer, *The Politics and Incentives of First Amendment Coverage*, 56 WM. & MARY L. REV. 1613 (2015).

environments work.² Each operates by reference to a familiar ideal of competition and contest according to which, as Justice Holmes put it, “the best test of truth is the power of the thought to get itself accepted in the competition of the market.”³ More importantly for purposes of this Essay, each focuses attention on particular kinds of market failure and suggests correspondingly particular criteria for market success, and each assumes certain structural preconditions within which market dynamics unfold.

A. From Distribution Bottlenecks to Microtargeting at Scale

One frame conventionally employed in evaluating speech regulation is the idea of the *distribution bottleneck*. A distribution bottleneck confers market power on whoever controls it, but the frame of the distribution bottleneck is not concerned with market power in the abstract. It is an artifact of mid-twentieth-century media regulation and litigation, and so it is also an artifact of the principal risk to free expression that mid-twentieth-century media technologies were thought to create: centralized, practically and technically unavoidable control of access to communication channels resulting in preemptive censorship. In the late twentieth century, as media technologies evolved and the power of media ownership began to manifest in ways that did not align with the frame, interested actors mobilized the frame to mount successful campaigns for deregulation. In the Internet era, platforms exercise power in ways that do not appear within the frame at all.

The distribution bottleneck frame originated in disputes about the constitutionality of regulations designed to limit the power of mass media owners. So, for example, because the then-usable broadcast spectrum imposed a natural bottleneck effect, the Federal Communications Commission imposed an access mandate—the fairness doctrine—on broadcast licensees to ensure that those wishing to express opposing or minority viewpoints had opportunities to respond to certain kinds of statements.⁴ In practice, the doctrine proved controversial, opening new vistas for gamesmanship and

² Frames, thus, are not analogies but rather sets of background assumptions about environmental structure and operation within which analogies sit. Cf. Heather Whitney, *Search Engines, Social Media, and the Editorial Analogy*, KNIGHT FIRST AMEND. INST. (Feb 27, 2018), <https://knightcolumbia.org/content/search-engines-social-media-and-editorial-analogy> [<https://perma.cc/56NX-X7ET>].

³ *Abrams v. United States*, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting).

⁴ See *Editorializing by Broadcast Licensees*, Report of the Commission, 13 F.C.C. 1246, 1249–50 ¶¶ 6–7 (1949) (implementing the fairness doctrine); *Applicability of the Fairness Doctrine in the Handling of Controversial Issues of Public Importance*, 29 Fed. Reg. 10422–23 (1964) (reaffirming that the fairness doctrine applies to television broadcast licensees); *Red Lion Broad., Co. v. FCC*, 395 U.S. 367, 400–01 (1969) (upholding the fairness doctrine against a facial challenge).

threatening to embroil regulators directly in content disputes, and the FCC ultimately withdrew it.⁵ Even after the fairness doctrine's demise, however, the distribution bottleneck frame survived. As technologies for cable and satellite distribution introduced multiple alternative channels for reaching viewers, courts began to rely on the frame to strike down new types of media regulation, reasoning that market pressures would ensure adequate alternative avenues of communicative opportunity. In response to such decisions, regulators gradually learned to color within the boundaries that the bottleneck frame imposed.⁶

The distribution bottleneck frame informs the modern landscape of anti-electioneering jurisprudence in two ways. First, the Court's evolving stance on the constitutionality of limits on campaign contributions reflects an analogous understanding of the relationship between scale and control. The earliest decisions upholding contribution limits painted large contributions as inevitably leading to corruption of the democratic process because they engendered a "pay-to-play" norm.⁷ According to that way of thinking, money deployed at scale functions in the manner of a bottleneck limiting access to political influence. It crowds out disfavored inputs to political decision-making, and it does so explicitly in the service of particular outcomes. Later decisions reversing course on the constitutionality of limits on independent expenditures characterized democratic politics as inherently transactional, observing that "[a]ll speakers, including individuals and the media, use money

⁵ See *Inquiry into Alternatives to the General Fairness Obligations of Broadcast Licensees*, 102 F.C.C.2d 143, 147–48, 246 (1985) (concluding "that the fairness doctrine, as a matter of policy, disserves the public interest..."); *Syracuse Peace Council v. FCC*, 867 F.2d 654, 669 (D.C. Cir. 1989) (upholding the FCC's order abolishing the fairness doctrine).

⁶ See *Reno v. ACLU*, 521 U.S. 844, 853 (1997) (observing that "[n]o single organization controls any membership in the Web, nor is there any single centralized point from which individual Web sites or services can be blocked from the Web."); *Denver Area Educ. Telecomms. Consortium v. FCC*, 518 U.S. 727, 776–78 (1996) (Souter, J., concurring) (discussing the difficulty of ruling on the constitutionality of Congress's permissive grant of authority to cable operators to regulate the content of leased independent programmers, given the "technological and regulatory flux" of industries in which its component "individual entities [can] act as bottlenecks to the free flow of information."); *Turner Broad. Sys., Inc. v. FCC*, 512 U.S. 622, 663 (1994) (upholding Sections 4 and 5 of the Cable Television Consumer Protection and Competition Act of 1992 while observing that the "First Amendment's command that government not impede the freedom of speech does not disable the government from taking steps to ensure that private interests not restrict, through physical control of a critical pathway of communication, the free flow of information and ideas"); Ellen P. Goodman, *Media Policy and Free Speech: The First Amendment at War With Itself*, 35 HOFSTRA L. REV. 1211, 1226–27 (2007).

⁷ See, e.g., *McConnell v. FEC*, 540 U.S. 93 (2003); *Nixon v. Shrink Missouri Gov't PAC*, 528 U.S. 377 (2000); *Buckley v. Valeo*, 424 U.S. 1 (1976).

amassed from the economic marketplace to fund their speech.”⁸ From that perspective, the bottleneck effect disappears. Campaign finance arrangements merely express the “power of the thought to get itself accepted in the competition of the market” and thereby mirror the Darwinian struggle for supremacy that Justice Holmes envisioned.⁹

Second, the bottleneck frame helps to explain how modern regimes of disclosure-based election regulation have chosen to handle the problem of anonymous speech. After *McIntyre v. Ohio Election Commission*,¹⁰ in which the Court invalidated a state prohibition on anonymous leafletting, many states amended their laws. They did not eliminate speaker identification requirements, but rather crafted narrow exceptions permitting anonymous election-related speech by relatively small-scale speakers.¹¹ Through the lens of strict scrutiny—or even that of “exacting scrutiny” as articulated in the Court’s later election regulation cases—that resolution is difficult to understand.¹² The concerns articulated by the *McIntyre* majority about chilling effects and failure to differentiate the messenger from the message apply equally to large, popular, and well-resourced actors.¹³ If statutory shelters for electioneering speech by small-scale speakers make sense at all, that can only be because (largely implicit) preconceptions about the necessary correlation between size and electoral influence render the state interest in disclosure about small-scale interventions much less compelling.

Platforms that combine networked economies of scale with capabilities for data-driven, algorithmic microtargeting and socially networked, cascading flows of information restructure the relationships between money, scale, and the possibility of improper influence in ways that defy earlier assumptions. Begin with bottleneck control. Arguably, today’s dominant advertiser-funded platforms qualify as distribution bottlenecks for content, but in other respects the analogy to the types of control enjoyed by mid-twentieth-century

⁸ *Citizens United v. FEC*, 558 U.S. 310, 351 (2009) (rejecting the antidistortion rationale adopted in *Austin v. Michigan Chamber of Commerce*, 494 U.S. 652 (1990)).

⁹ *Abrams v. United States*, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting); Vincent Blasi, *Holmes and the Marketplace of Ideas*, 2004 SUP. CT. REV. 1, 24–33 (2004). I am indebted to Neil Richards for this point.

¹⁰ 514 U.S. 334 (1995).

¹¹ See, e.g., Del. Code Ann., Tit. 15, § 8021 (2020); Fla. Stat. § 106.143 (2020); Ky. Rev. Stat. Ann. § 121.190 (2020); La. Rev. Stat. Ann. § 18:1463 (2020); Me. Rev. Stat. Ann. Tit. 21-A, § 1014 (2020); N.D. Cent. Code § 16.1-10-04.1 (2020); S.C. Code Ann. § 8-13-1354 (2020); Tenn. Code Ann. § 2-19-120 (2020); Tex. Elec. Code Ann. § 255.001 (2020); W. Va. Code § 3-8-12(a) (2020); Wyo. Stat. § 22-25-110 (2020).

¹² Exacting scrutiny requires disclaimer and disclosure requirements to bear a “substantial relation” to a “sufficiently important” governmental interest. See *Citizens United*, 558 U.S. at 366–67; *McConnell*, 540 U.S. at 231–32; *Buckley*, 424 U.S. at 64–66; see also R. George Wright, *A Hard Look at Exacting Scrutiny*, 85 UMKC L. REV. 207, 209–11 (2016).

¹³ *McIntyre*, 514 U.S. at 357.

television and radio networks is difficult to sustain. Platforms like Google, Facebook, and Twitter have thrived precisely because they enable certain forms of access to the so-called long tail—i.e., content of interest to only a small number of readers. As long as they are willing to use the dominant platforms' services, thereby foregoing direct access to both user data and information about algorithm design and training, would-be speakers of all sizes and persuasions can buy targeted advertising at a relatively low cost. Platforms might exercise preemptive control of the *content* of such ads, but for the most part they do not, as that would be much less profitable. Their interest lies simply in extracting surplus from whatever types of messaging elicit responses (positive or negative, but in any event data-generating) from their customers.¹⁴

Even so, the combination of platforms' own economic self-interest and the narrower interests of those who purchase and compete for digital advertising reshapes the universe of information available to users. Platforms win when they can promise the most comprehensive and accurate methods of targeting content based on predicted interest and the largest pool of potential viewers of that content; advertisers win when they achieve clickthrough, and content providers win when they can promise advertisers higher pageviews via either targeting or social sharing of their content. For platforms, competition for eyeballs both incentivizes and rewards interface design that keeps users on the platform and tracks them carefully and comprehensively as they browse, click, like, hate, comment on, and share items with one another.¹⁵ For advertisers and content providers, competition for eyeballs both incentivizes and rewards content design for maximal "engagement" as defined by those activities.¹⁶ The resulting effects have been termed "filter bubbles,"

¹⁴ See JULIE E. COHEN, *BETWEEN TRUTH AND POWER: THE LEGAL CONSTRUCTIONS OF INFORMATIONAL CAPITALISM* 63–86 (2019); SHOSHANA ZUBOFF, *THE AGE OF SURVEILLANCE CAPITALISM: THE FIGHT FOR A HUMAN FUTURE AT THE NEW FRONTIER OF POWER* 63–97 (2019); Ryan Calo, *Digital Market Manipulation*, 82 GEO. WASH. L. REV. 995 (2014). Public outrage about some kinds of unambiguously illegal content has produced a few high-profile exceptions. See, e.g., T.J. McIntyre, *Child Abuse Images and Cleanfeeds: Assessing Internet Blocking Systems*, in RESEARCH HANDBOOK ON GOVERNANCE OF THE INTERNET 277, 277–95 (Ian Brown ed., 2012); Jack Gillum & Ariana Tobin, *Facebook Won't Let Employers, Landlords or Lenders Discriminate in Ads Anymore*, PROPUBLICA (Mar. 19, 2019, 2:00 PM), <https://www.propublica.org/article/facebook-ads-discrimination-settlement-housing-employment-credit> [<https://perma.cc/9RSJ-QQHE>].

¹⁵ See, e.g., ADAM ALTER, *IRRESISTIBLE: THE RISE OF ADDICTIVE TECHNOLOGY AND THE BUSINESS OF KEEPING US HOOKED* (2017); ZUBOFF, *supra* note 14, at 159–62, 457–74; Tristan Harris, *The Slot Machine in Your Pocket*, DER SPIEGEL ONLINE (July 27, 2016, 5:25 PM), <http://www.spiegel.de/international/zeitgeist/smartphone-addiction-is-part-of-the-design-a-1104237.html> [<https://perma.cc/V3AA-ZBWM>].

¹⁶ See FRANKLIN FOER, *WORLD WITHOUT MIND: THE EXISTENTIAL THREAT OF BIG TECH*

but that term is to some extent misleading. Platform users do not experience or self-select into impermeable bubbles but rather sort themselves into opposing tribes. They respond most readily and predictably to content that reinforces their tribal inclinations—especially content that triggers outrage or affords opportunities to signal affiliation—and they search for content using syntax that prompts algorithms to serve up tribally validating results.¹⁷

Market dominance plays a role in this story—platforms win most decisively when they can promise the largest pools of potential viewers for any and all content—but disrupting the dominance of any particular platform would not cure the dysfunctions that more widely distributed capabilities for personalization at scale and optimization for engagement now create. In a networked media ecosystem designed for content targeting, optimization for engagement, and amplification of social flows, polarized and polarizing content spreads rapidly from one platform to another and between online and traditional media, gaining in volume as it travels.¹⁸

Under such conditions, the implicit presumption about the relative inefficacy of small-scale interventions also no longer holds. Because information flows within platform-based, massively intermediated environments are data-driven and social, provocations from the margins can be designed to trigger patterns of rapid, cascading spread. Such provocations exploit properties of human behavior—most notably, fear of missing out on what everyone else already knows; properties of social networks—particularly

131–54 (2017); Caitlin Petre, *The Traffic Factories: Metrics at Chartbeat, Gawker Media, and The New York Times*, TOW CTR. FOR DIG. JOURNALISM (2015), <https://doi.org/10.7916/D80293W1> [<https://perma.cc/4L4V-5GLU>]; Antonio Garcia Martinez, *How Trump Conquered Facebook—Without Russian Ads*, WIRED (Feb. 23, 2018, 10:06 AM), <https://www.wired.com/story/how-trump-conquered-facebookwithout-russian-ads/> [<https://perma.cc/9APZ-HPZ7>] (“Facebook uses a complex model that considers both the dollar value of each bid as well as how good a piece of clickbait (or view-bait, or comment-bait) the corresponding ad is. If Facebook’s model thinks your ad is 10 times more likely to engage a user than another company’s ad, then your effective bid at auction is considered 10 times higher than a company willing to pay the same dollar amount.”); see also Kerry Jones, Kelsey Libert, & Kristin Tynski, *The Emotional Combinations that Make Stories Go Viral*, HARV. BUS. REV. (May 23, 2016), <https://hbr.org/2016/05/research-the-link-between-feeling-in-control-and-viral-content> [<https://perma.cc/VZ6V-H7WT>].

¹⁷ See Alice E. Marwick, *Why Do People Share Fake News? A Sociotechnical Model of Media Effects*, 2 GEO. L. TECH. REV. 474 (2018); Francesca Tripodi, *Searching for Alternative Facts: Analyzing Scriptural Inference in Conservative News Practices*, DATA & SOC’Y (May 16, 2018), <https://datasociety.net/library/searching-for-alternative-facts/> [<https://perma.cc/2QKK-3NYK>]. The “filter bubble” terminology originated with ELI PARISER, *THE FILTER BUBBLE: WHAT THE INTERNET IS HIDING FROM YOU* (2011).

¹⁸ On the interconnectedness of online and traditional media, see generally Ulrike Klinger & Jakob Svensson, *The Emergence of Network Media Logic in Political Communication: A Theoretical Approach*, 17 NEW MEDIA & SOC’Y 1241 (2015). See also Erin C. Carroll, *News as Surveillance*, 59 WASHBURN L. J. (forthcoming 2020).

their hub-and-spoke organization, which permits rapid spread via well-connected nodes; and properties of organizational behavior—especially traditional media outlets’ eagerness to chase and report on topics trending online.¹⁹ Multiple teams of researchers studying election manipulation have mapped the resulting patterns, tracing the paths followed by extreme and inflammatory content as it migrates from the periphery to the center of public consciousness.²⁰ Some such interventions originate with well-resourced state actors and powerful domestic political blocs, but others have been true bottom-up efforts.²¹

Because the distribution bottleneck frame originated in a world characterized by hierarchical control of content prepared for distribution to mass audiences, it has little of direct significance to say about either the operation or the distinctive dysfunctions of platform-based, massively intermediated information environments. It comprehends neither the sorts of personalized microtargeting that platform-based information infrastructures enable nor the ways that optimization for data-driven surplus extraction and competition for eyeballs incentivize self-sorting into political tribes hardened in their mutual contempt for one another. Yet it constitutes an imagined world in which the very possibility that third parties might hijack and weaponize socially networked flows is already foreclosed. Legislators attempting to craft new anti-electioneering laws for the platform era and courts reviewing such efforts should recognize that the distribution bottleneck frame has no place in either exercise.

¹⁹ See ALBERT-LÁSZLÓ BARABÁSI, *NETWORK SCIENCE* (2016) (properties of social networks); CASS R. SUNSTEIN, *ON RUMORS: HOW FALSEHOODS SPREAD, WHY WE BELIEVE THEM, AND WHAT CAN BE DONE* (2014) (information cascades); Marina Milyavskaya, et al., *Fear of Missing Out: Prevalence, Dynamics, and Consequences of Experiencing FOMO*, 42 *MOTIVATION & EMOTION* 725 (2018) (fear of missing out); Joan Donovan, *Source Hacking: Media Manipulation in Practice*, *DATA & SOC’Y* (Sept. 4, 2019), <https://datasociety.net/library/source-hacking-media-manipulation-in-practice/> [https://perma.cc/4DM7-8HR8] (traditional media outlets).

²⁰ See, e.g., Samantha Bradshaw & Philip N. Howard, *The Global Disinformation Order: 2019 Global Inventory of Organised Social Media Manipulation*, OXFORD PROJECT ON COMPUTATIONAL PROPAGANDA WORKING PAPER 2019.3 (2019), <https://comprop.oii.ox.ac.uk/research/cybertroops2019/> [https://perma.cc/2U67-G8UY]; Matthew Hindman & Vlad Barash, *Disinformation, ‘Fake News’ and Influence Campaigns on Twitter*, THE KNIGHT FOUND. (Oct. 2018), https://kf-site-production.s3.amazonaws.com/media_elements/files/000/000/238/original/KF-DisinformationReport-final2.pdf [https://perma.cc/8FQQ-NPRV].

²¹ See Vivian Ho, *The California Senator Fighting for the Strictest Vaccination Laws in the US*, THE GUARDIAN (Aug. 29, 2019, 2:00 PM), <https://www.theguardian.com/us-news/2019/aug/28/richard-pan-california-vaccines-strictest-law> [https://perma.cc/925G-5M92]; Whitney Phillips, *The Oxygen of Amplification: Better Practices for Reporting on Extremists, Antagonists, and Manipulators*, *DATA & SOC’Y* (May 22, 2018), <https://datasociety.net/library/oxygen-of-amplification/> [https://perma.cc/S8HZ-BKEM].

B. From Autonomy to Automaticity

A second frame conventionally employed in designing and evaluating speech regulation is the idea of the *rational listener*. The rational listener tests ideas for their persuasiveness and vets factual propositions for their truthfulness but has no interest in efforts to impose general, *ex ante* restrictions on the flow of low-quality ideas and propositions. The rational listener is autonomous and perspicacious and therefore (largely) self-reliant, capable of separating fact from falsehood and reason from self-interested conniving and demagoguery.²² Legislative design for the rational listener accordingly emphasizes transparency and informed choice, and courts have tended to regard such approaches as acceptable ways of advancing state interests precisely because they leave room for rational listeners to make their own decisions.

The idea of the rational listener has deep roots in the Anglo-American political tradition and more direct and immediate roots in the American system of political economy. Its first judicial articulation emerged in early twentieth-century cases involving restrictions on political liberty.²³ The contemporary rational listener frame, however, is also and importantly an artifact of mid-twentieth-century litigation over economic and consumer protection regulation. Thus, it is also an artifact of the particular risks that mid-twentieth-century consumer markets were thought to create—risks involving the emergence of more complex consumer products and services that consumers themselves could not easily evaluate. Regulators responded to those developments by prohibiting certain kinds of deception and requiring certain kinds of disclosure. The rational listener frame, however, dictated that ultimate decision-making authority should remain with the individual to the greatest extent practicable. Courts therefore struck down laws regulating advertising, labeling, and similar matters that seemed to be attempts to superimpose government judgments about the ultimate desirability of the covered products and services.²⁴ To legislators and regulators, such decisions communicated a

²² See Lyrixa Barnett Lidsky, *Nobody's Fools: The Rational Audience as First Amendment Ideal*, 2010 U. ILL. L. REV. 799 (2010).

²³ See *Whitney v. California*, 274 U.S. 357, 372 (1927) (Brandeis, J., concurring); *Schenck v. United States*, 249 U.S. 47 (1919); *Abrams v. United States*, 250 U.S. 616, 624 (1919) (Holmes, J. dissenting). See generally G. Edward White, *The First Amendment Comes of Age: The Emergence of Free Speech in Twentieth-Century America*, 95 MICH. L. REV. 299 (1996).

²⁴ See, e.g., *Thompson v. W. States Med. Ctr.*, 535 U.S. 357 (2002); *Lorillard Tobacco Co. v. Reilly*, 533 U.S. 525 (2001); *Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, Inc.*, 425 U.S. 748 (1976); *Nat'l Association of Manufacturers v. SEC*, 800 F.3d 518 (D.C. Cir. 2015). *But see Florida Bar v. Went for It, Inc.*, 515 U.S. 618 (1995); *United States v. Edge Broad. Co.*, 509 U.S. 418 (1993). Emergent consumer complexity continues to test

clear preference for laws and regulations that focused simply on injecting more or different kinds of information into the marketplace and into public discourse.

Within the modern landscape of anti-electioneering jurisprudence, the rational listener frame is especially prominent in disputes about laws mandating disclosure of campaign contributions and advertising expenditures. In an era when the anti-corruption rationale for upholding spending limits no longer holds sway, federal election regulation depends ever more heavily on such provisions. Litigation over their constitutionality has given rise to the idea of an “informational interest” that is sufficiently important to warrant (slightly) relaxed scrutiny—“exacting” rather than “strict”—and also to override anonymity interests in certain circumstances.²⁵ Over time, the Court’s opinions elaborating the informational interest have leaned heavily on the frame of the rational listener. In particular, within the more recently developed conception of purchased access as consistent with a broadly transactional democratic politics, disclosure “enables the electorate to make informed decisions and give proper weight to different speakers and messages.”²⁶

Many current reform proposals for tackling platform-based disinformation double down on transparency, proposing to require disclosures about a variety of matters including ad buys, ad targeting, and automated “bot” speech.²⁷ In the abstract, such proposals sound like great ideas. One might even hope for new, technologically mediated advances in electoral transparency. Networked information technologies have already facilitated widespread, easy access to data about contributions to political campaigns; now, machine learning techniques can be trained on other categories of disclosed data to map networks of influence.²⁸ Rational listeners who discover that they have been consorting with bots can reevaluate their choices.

the outer limits of the doctrinal commitment to protecting the rational listener. *Cf.* David C. Vladeck, *The Difficult Case of Direct-To-Consumer Drug Advertising*, 41 *LOY. L.A. L. REV.* 259 (2007).

²⁵ See *Citizens United*, 558 U.S. at 369; *Buckley*, 424 U.S. at 81; *see, e.g.*, Daniel R. Ortiz, *The Informational Interest*, 27 *J.L. & POL.* 663 (2012); Lear Jiang, Note, *Disclosure’s Last Stand? The Need to Clarify the “Informational Interest” Advanced by Campaign Finance Disclosure*, 119 *COLUM. L. REV.* 487 (2019).

²⁶ *Citizens United*, 558 U.S. at 371.

²⁷ *See, e.g.*, Bot Disclosure and Accountability Act of 2019, S. 2125, 116th Cong. (2019); Honest Ads Act, S. 1989, 115th Cong. (2017); The B.O.T. (Bolstering Online Transparency) Act, S. 1001, 2017-2018 Leg., Reg. Sess. (Cal. 2018) (codified in Cal. Bus. & Prof. Code § 17941(a)).

²⁸ For an especially thoughtful take on what systemic transparency requirements might reveal, see Ellen P. Goodman, *Digital Information Fidelity and Friction*, *KNIGHT FIRST AMEND. INST.* (Feb. 26, 2020) <https://knightcolumbia.org/content/digital-fidelity-and-friction> [<https://perma.cc/BNR3-9APR>]; *see also* Richard L. Hasen, *Deep Fakes, Bots, and Siloed*

Optimism about the potential for greater disclosure to ameliorate the problems caused by microtargeting is misplaced, however, because platform-based, massively intermediated information environments are not designed for the rational listener. Instead, they are both systematically configured and continually reoptimized to elicit automatic, precognitive interactions with online content. As noted in Section A, the currency of the platform-based environment is user behavioral data, and that reality dictates a set of interrelated strategies for platform providers. Platform interfaces work to normalize consent to tracking; to keep users on the platform to facilitate the most comprehensive tracking; and to harvest data about user preferences and aversions using low-level stimulus-response feedback loops—e.g., “buttons” for liking and sharing content—designed for automatic, habitual engagement.²⁹ Behind the scenes, platform algorithms work to derive behavioral and psychographic profiles based on user engagement data; to drive socially-networked flows of content; and to amplify such flows in ways that maximize advertising revenues.³⁰ And, as noted above, these characteristics of platform-based environments engender complementary strategies for advertisers and content providers, who work to design “clickbait” and foster its widest possible circulation.

Skeptics and advocates seeking to minimize alarm about the effects of platform capabilities for voter microtargeting argue that microtargeting promises more than it delivers because it cannot change minds, but that

Justices: American Election Law in a Post-Truth World, 64 ST. LOUIS U. L.J. (forthcoming 2020). For useful introductions to some of the relevant methods; see Hindman & Barash, *supra* note 20; *Analytics & Predictive Models for Social Media*, TUTORIAL AT THE WORLD WIDE WEB CONFERENCE (Mar. 29, 2011), <http://snap.stanford.edu/proj/socmedia-www/socMedia-www11-part1.pdf> [<https://perma.cc/YGA9-9REU>].

²⁹ On normalizing consent and the role of “dark patterns,” see generally WOODROW HARTZOG, *PRIVACY’S BLUEPRINT: THE BATTLE TO CONTROL THE DESIGN OF NEW TECHNOLOGIES* (2018); Arunesh Mathur et al., *Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites*, in 22ND ACM CONFERENCE ON COMPUTER-SUPPORTED COOPERATIVE WORK AND SOCIAL ENGINEERING 81 (2019); Jamie Luguri & Lior Strahilevitz, *Shining a Light on Dark Patterns* (U. Chicago, Poverty Law Working Paper No. 719, Aug. 7, 2019). On principles of addictive design and their application to digital interfaces, see ALTER, *supra* note 15; Kyle Langvardt, *Regulating Habit-Forming Technology*, 88 FORDHAM L. REV. 129 (2019); see also NATASHA DOW SCHÜLL, *ADDICTION BY DESIGN: MACHINE GAMBLING IN LAS VEGAS* (2012). On the development of the like and share buttons as mechanisms for data harvesting, see ZUBOFF, *supra* note 14, at 457–60.

³⁰ For the foundational research on psychographic profiling, see Michal Kosinski, David Stillwell, & Thore Graepel, *Private Traits and Attributes Are Predictable from Digital Records of Human Behavior*, 110 PNAS 5802 (2013); Wu Youyou, Michal Kosinski, & David Stillwell, *Computer-Based Personality Judgments Are More Accurate than Those Made by Humans*, 112 PNAS 1036 (2015). On current platform implementations, see ZUBOFF, *supra* note 14 at 270–92; Nabihah Syed, *Real Talk about Fake News: Towards a Better Theory for Platform Governance*, 127 YALE L.J. F. (Oct. 9, 2017).

argument mistakes the purposes for which microtargeting is more commonly deployed by political operatives. Without a doubt, inducing undecided (or actively hostile) voters to vote for a particular candidate is much harder than inducing them to order of-the-moment, celebrity-endorsed sneakers or book an ostensibly discounted stay at a luxury hotel (“only three rooms left at this price!”). Using behavioral and psychometric targeting techniques to play on recipients’ fears and to activate their tribal loyalties and enmities, however, is a different—and much easier—proposition.

The result of platform design for maximal data harvesting, continual user engagement, and cascading, socially networked spread based on automatic, conditioned responses is a networked digital environment in which the rational listener’s presumptive autonomy increasingly is displaced by automaticity—by habitual, precognitive behaviors that require no conscious attention. Platform-based environments constitute what legal philosopher Mireille Hildebrandt terms the digital unconscious, a field of operation within which agency is mindless, data-driven and characterized by “ubiquitous anticipation” of user predispositions.³¹ The individual subject of the digital unconscious is not the rational listener but rather the listener who is not really listening at all. Critically, moreover, the digital unconscious is also a “a field of operation for precognitive activation and manipulation at scale.”³² Voter microtargeting efforts move and are designed to move on the collective level, nurturing rumor and innuendo, hardening targeted populations in their tribal responses to real and perceived differences, and frustrating the sorts of efforts toward rapprochement on which theories about republican self-government rely.

The rational listener frame, which foregrounds the autonomous individual, cannot make sense of the platform-based information environment. Regulatory initiatives based on mandated disclosure, which are oriented toward the needs and presumed competencies of the rational listener, fatally misapprehend platforms’ operative logics and scalar effects. So too with solutions based on fact-checking by third parties, whose interventions must battle upstream against an unrelenting torrent of bias reinforcement, and those based on opt-out rights, which rely on recipients themselves to recognize and disavow their own most automatic and deeply-ingrained habits and affinities. Legislators attempting to craft new anti-electioneering laws for the platform

³¹ MIREILLE HILDEBRANDT, SMART TECHNOLOGIES AND THE END(S) OF LAW 66–67 (2015).

³² Julie E. Cohen, *The Emergent Limbic Media System*, in LIFE AND THE LAW IN THE ERA OF DATA-DRIVEN AGENCY 60, 61 (Mireille Hildebrandt & Kieron O’Hara, eds. 2019). See ZUBOFF, *supra* note 14 at 295–99 (describing population-based strategies of tuning, herding, and conditioning); see also Madeline Lamo & Ryan Calo, *Regulating Bot Speech*, 66 UCLA L. REV. 988, 1017 (2019). Cf. Helen Norton, *Powerful Speakers and Their Listeners*, 90 U. COLO. L. REV. 441 (2019).

era and courts reviewing such efforts should understand that the rational listener cannot help them.

C. From “Neutral Tools” to Amplified Flows

The third frame conventionally employed in evaluating speech regulation—one that comes into play when the legal responsibility of third-party intermediaries is at stake—is the idea of the *intentional facilitator*. According to this frame, a third-party intermediary should not automatically incur liability for harms caused by information circulated by others. For both legislators and courts, that prospect raises worries about censorship by proxy. The intentional facilitator frame counteracts those worries by linking liability to some type of knowing involvement with specific items or categories of clearly illegal content.

The intentionality frame powerfully infuses two very different statutes governing responsibility for online content that are widely understood as encoding opposite policy choices. One is Section 230 of the Communications Decency Act, which immunizes interactive service providers from liability for unlawful content published by users of their services unless they have played a role in its development.³³ The other, Section 512 of the Copyright Act, creates safe harbors for information intermediaries but withholds safe harbor from intermediaries that have failed to act upon receiving knowledge of specific infringing content or that have offered services specifically designed to profit from infringing flows.³⁴ Intentionality plays a central role in both regimes even though the conditions for loss of immunity differ. In particular, although Section 230’s drafters sought to limit the effect of background doctrines tying liability to mere knowledge, and Section 230’s contemporary defenders regard the Copyright Act’s notice-and-takedown regime as antithetical to Section 230’s animating spirit, both statutes link liability for facilitating the spread of harmful content to volitional involvement with specific content rather than to the underlying design of distribution mechanisms more generally.³⁵

³³ 47 U.S.C. § 230 (2012).

³⁴ 17 U.S.C. § 512(c)-(d) (2012).

³⁵ On the intent behind Section 230, see H.R. Rep. No. 104-458, at 190 (1996) (“The conferees intend that [CDA Section 230’s] defense be construed broadly to avoid impairing the growth of online communications through a regime of vicarious liability.”); see also *id.* at 194; Robert Cannon, *The Legislative History of Senator Exon’s Communications Decency Act: Regulating Barbarians on the Information Superhighway*, 49 FED. COMM. L.J. 51, 61–70 (1996). For representative contemporary reactions to the prospect of replacing Section 230 with a notice-and-takedown regime, see DANIELLE CITRON, HATE CRIMES IN CYBERSPACE 171–89 (2014); Mike Masnick, *Thanks To Copyright, We Already Know How Aggressive Content Moderation*

Notably, courts evaluating disputes under both Section 230 and Section 512 have homed in on the underlying commonality, framing networked digital technologies—including platform-based, massively intermediated information environments—as neutral tools that are, and should be, exempt from more intrusive oversight. So, for example, in cases about the scope of Section 230’s immunity, courts have opined that an online dating service that failed to implement certain safety features could not be penalized because it simply “provid[ed] ‘neutral assistance’ in the form of tools and functionality available equally to bad actors and . . . intended users”³⁶; and that a roommate matching service could not be responsible for discriminatory requests posted in the spaces it provided for unstructured comments because liability for providing a “simple, generic prompt” would be inconsistent with the immunity afforded to services “that provide users *neutral* tools to post content online.”³⁷ In cases about the scope of Section 512’s safe harbors, they have rejected interpretations that would impose liability based on general awareness of likely infringement because such interpretations would require platforms to monitor their systems for signs of illegality.³⁸

In the abstract, there are sound policy reasons for worrying about the effects of liability for tool developers. In particular, courts and commentators worry with good reason that takedown obligations could morph into an open-ended mandate to sanitize the universe of public information by removing controversial content.³⁹ Concerns about giving copyright interests de facto control over technological development also are well taken.⁴⁰ It is worth noting, though, that the quasi-religious devotion to untrammelled innovation that sometimes accompanies such concerns is both an historical anomaly and

Works: And It’s A Disaster, TECHDIRT (Oct. 4, 2018, 9:13 AM), <https://www.techdirt.com/articles/20180927/13030040731/thanks-to-copyright-we-already-know-how-aggressive-content-moderation-works-disaster.shtml> [https://perma.cc/4PH4-53GR].

³⁶ *Herrick v. Grindr LLC*, 765 Fed. App’x 586, 591 (2d. Cir.), *cert. denied*, 140 S. Ct. 221 (2019).

³⁷ *Fair Hous. Council of San Fernando Valley v. Roommates.com, LLC*, 521 F.3d 1157, 1174–75 (9th Cir. 2008) (en banc).

³⁸ *UMG Recordings, Inc. v. Shelter Capital Partners LLC*, 718 F.3d 1006, 1022–23 (9th Cir. 2013); *Viacom Intern., Inc. v. YouTube, Inc.*, 676 F.3d 19, 35 (2d. Cir. 2012).

³⁹ *See, e.g.,* Joris van Hoboken and Daphne Keller, *Design Principles for Intermediary Liability Laws*, 4, TRANSATLANTIC WORKING GRP. ON CONTENT MODERATION ONLINE & FREEDOM OF EXPRESSION (Oct. 8, 2019), https://www.ivir.nl/publicaties/download/Intermediary_liability_Oct_2019.pdf [https://perma.cc/B9SJ-M7Y5]; Danielle Keats Citron, *Extremist Speech, Compelled Conformity, and Censorship Creep*, 93 NOTRE DAME L. REV. 1035 (2018).

⁴⁰ *See, e.g.,* Anupam Chander, *How Law Made Silicon Valley*, 63 EMORY L.J. 639, 663 (2014); Mark A. Lemley & R. Anthony Reese, *Reducing Digital Copyright Infringement Without Restricting Innovation*, 56 STAN. L. REV. 1345, 1356 (2004).

an anti-regulatory dog whistle. Major sectors of the modern regulatory state emerged precisely to constrain innovation's excesses, and although the design of regulatory oversight mechanisms has engendered profound disagreements, support for such basic propositions as, say, the continued existence of agencies devoted to environmental protection and food and drug oversight is broad and durable.

The "neutral tools" characterization, however, is overly simplistic both in general and as applied to platform-based, massively intermediated information environments. Generally speaking, tools reflect the priorities of their designers and may disserve or simply overlook other priorities and needs.⁴¹ We have already seen that the platform-based, massively intermediated environment alters the universe of information in specific, non-neutral ways. Platforms' formal agnosticism about information *content* belies an operational orientation that reliably infuses information flows with distinctive attributes. The relatively crude distinction between knowing involvement and the mere provision of neutral tools for accessing information elides design principles that privilege polarization, amplification, and automaticity, and those principles shape both the content and the consumption of networked, massively intermediated communication.

In the context of platform-based, massively intermediated environments, the legal system should be less concerned with intentionality as to specific pieces of content—a lens that inevitably implicates the state in choice of political preferences—and more concerned with a deliberate design orientation that privileges automatic, habitual response and reflexive amplification. As currently constituted, the platform-based, massively intermediated information environment is an arena for Darwinian struggle in which the determinant of superiority is not truth but rather bias confirmation. The First Amendment does not require legislators or judges to privilege design for automaticity and reflexive amplification, and it permits them to select a frame that makes such choices and their undeniable, empirically demonstrated effects more salient.

⁴¹ For a good general introduction to the social construction of tools and technologies, see Wiebe E. Bijker, *OF BICYCLES, BAKELITES, AND BULBS: TOWARD A THEORY OF SOCIOTECHNICAL CHANGE* (1995). Cf. Daphne Keller, *Toward a Clearer Conversation About Platform Liability*, KNIGHT FIRST AMEND. INST. (April 6, 2018), <https://knightcolumbia.org/content/toward-clearer-conversation-about-platform-liability> [<https://perma.cc/E79Q-6PDS>] ("All of this makes neutrality something of a Rorschach test. It takes on different meanings depending on the values we prioritize.").

III. FROM INFORMATION MARKETPLACES TO INFORMATION PLATFORMS: FRAMING THE FIRST AMENDMENT'S FUTURE

Understanding the ways that platform-based, massively intermediated information environments work, and the ways that such environments engender unacceptable structural conditions for public discourse, suggests a new frame to be used in designing and evaluating speech regulation: that of the platform seen for itself. The interests implicated by this frame are not simply informational, and the compelling need to protect them justifies both new types of regulatory oversight and new ways of thinking about the associated tailoring problems.

It is useful to begin with definitions. As applied to networked information intermediaries, the term “platform” is a metaphor, one that has worked both to draw attention to certain features of platform-mediated spaces and deflect attention from others.⁴² Sustained scrutiny of information platforms, however, has surfaced more information about their attributes and capabilities, making it possible to describe those attributes and capabilities in more precise ways that could inform new framework legislation. An *information platform* is an information intermediary that uses data-driven, algorithmic methods and standardized, modular interconnection protocols to facilitate digitally networked interactions and transactions among its users. As that general definition is intended to suggest, a platform-based environment might be designed in a variety of ways. This Essay, however, has identified the following capabilities that have become characteristic of contemporary platform-based environments: collection of highly granular data about user behaviors; design of interfaces to elicit behavioral data via automatic, conditioned responses; processing of such data to create behavioral and psychometric profiles of users and user populations; targeting of content to users and user populations based on such profiles; and algorithmically-mediated amplification of content based on user engagement.

The definition articulated above also makes clear what platforms are not: they are not publishers, nor are they public fora as that concept has conventionally been understood and elaborated within First Amendment jurisprudence and theory.⁴³ Platforms are private, for-profit entities that

⁴² Tarleton Gillespie, *The Politics of 'Platforms,'* 12 NEW MEDIA & SOC'Y 3 (2010).

⁴³ For representative examples of such arguments, see Eric Goldman, *Section 230 Applies to Facebook's Post Removals and Account Suspensions—King v. Facebook*, TECH. & MKTG L. BLOG (Sept. 6, 2019), <https://blog.ericgoldman.org/archives/2019/09/section-230-applies-to-facebooks-post-removals-and-account-suspensions-king-v-facebook.htm> [<https://perma.cc/TDU7-3HYS>]; Mike Masnick, *Supreme Court Signals Loud and Clear That Social Media Sites Are not Public Forums That Have to Allow All Speech*, TECHDIRT (Jun.

operate as central nodes in the contemporary personal data economy. They afford their users opportunities for self-expression because self-expression generates behavioral data that can be monetized. They route content (or, more accurately, links to content published by others) using predictive algorithms that have been trained on user behavioral data, and they amplify socially networked flows in ways that elicit conditioned, automatic, and tribal responses because that is the approach that most reliably enriches their shareholders and venture investors. Seen for themselves, platforms merit neither the solicitude traditionally accorded publishers wishing to express their opinions nor the rote, unthinking application of rules traditionally applied to institutions performing public access functions.⁴⁴

Platform capabilities do not simply threaten the informational interest long recognized in the Court's election jurisprudence. They also threaten other interests that are important both instrumentally—i.e., as ways of ensuring fidelity to the informational interest—and intrinsically because they are inseparably intertwined with preservation of a system of government that is accountable both to the people it serves and to the rule of law. (To be clear, platform capabilities also implicate interests that I do not discuss here. For example, anti-vaxxer propaganda that risks undermining herd immunity jeopardizes an important interest in public health. For purposes of anti-electioneering regulation, however, the interests described below are key.)

The first interest threatened in platform-based, massively intermediated information environments is an *anti-factionalism interest*. As Anthony Johnstone has explained, such an interest is both latent in some strands of contemporary anti-electioneering jurisprudence and solidly grounded in an original understanding of the Constitution.⁴⁵ In *The Federalist* No. 10, Madison cautioned explicitly and pointedly against the threat posed by factions that might first capture and then subvert the institutions of democratic government by subordinating public functions to their own

18, 2019, 10:42 AM), <https://www.techdirt.com/articles/20190617/16001942415/supreme-court-signals-loud-clear-that-social-media-sites-are-not-public-forums-that-have-to-allow-all-speech.shtml> [<https://perma.cc/444W-VLJA>]; see also John Herrman, *How Hate Groups Forced Online Platforms to Reveal Their True Nature*, N.Y. TIMES MAG. (Aug. 21, 2017).

⁴⁴ I intend no comment on whether, having opened social media accounts, government officials must manage those accounts in a manner consonant with public forum doctrine. That question is both analytically distinct from questions about the status of platforms themselves and far more amenable to straightforward doctrinal analysis. See *Knight First Amend. Inst. v. Trump*, 928 F.3d 226 (2d Cir. 2019); *Davison v. Randall*, 912 F.3d 666 (4th Cir. 2019). But see *Morgan v. Bevin*, No. 3:17-CV-00060-GFVT-EBA, 2018 U.S. Dist. LEXIS 204657 (E.D. Ky. Dec. 3, 2018) (confusing the two questions).

⁴⁵ Anthony Johnstone, *A Madisonian Case for Disclosure*, 19 GEO. MASON L. REV. 413, 443–69 (2012).

narrower interests.⁴⁶ Modern arguments for translation of the anti-factionalism interest into the domain of election regulation have focused on reviving the broader, now-disfavored, understanding of campaign finance as part of a conscious return to civic republicanism.⁴⁷ But the anti-factionalism interest also bears on the ongoing debate about the structural properties of platform-based speech environments. The centrifugal properties of the platform-based environment—within which communications are systematically optimized to elicit, separate, and harden tribal reflexes—enable powerful factions to weaponize networked information flows in order to perpetuate their own power and advantage. At the same time, they disable the collective capacity to produce and propagate gap-bridging responses.

The second interest threatened in platform-based, massively intermediated information environments is an *anti-manipulation interest*. As defined by Daniel Susser, Beate Roessler, and Helen Nissenbaum, manipulation means hidden interference that deprives us of authorship over our own choices.⁴⁸ As Susser, Roessler, and Nissenbaum argue, if a rule against manipulation is to have any concrete force, it must apply to the structure of the networked communications environment rather than just to particular, discrete communications that contain manipulative content.⁴⁹ Manipulation in platform-based information environments is neither occasional nor accidental; it is endemic and results from capabilities that platforms systematically design, continually reoptimize, and deliberately offer up to third parties for exploitation. Properly conceived, the anti-manipulation interest encompasses the dark patterns that keep users enrolled and logged in, the stimulus-response loops designed to elicit automatic, precognitive responses and harvest the resulting data, and the mechanisms for harnessing that data to enable microtargeting based on user vulnerabilities and fears.

Finally, the emergent properties of information flows in platform-based, massively intermediated environments threaten a structural interest that warrants separate recognition. This interest, which I will call an *anti-authoritarianism interest*, concerns the stability and robustness of foundational democratic institutions and requires us to confront another underlying presumption of the marketplace-of-ideas metaphor that underwrites so much of First Amendment jurisprudence and theory. Implicit

⁴⁶ See THE FEDERALIST NO. 10 (James Madison).

⁴⁷ See, e.g., ZEPHYR TEACHOUT, CORRUPTION IN AMERICA: FROM BENJAMIN FRANKLIN'S SNUFF BOX TO CITIZENS UNITED (2016); LAWRENCE LESSIG, REPUBLIC, LOST: HOW MONEY CORRUPTS CONGRESS—AND A PLAN TO STOP IT (2011); see generally Lawrence Lessig, *Fidelity in Translation*, 71 TEX. L. REV. 1165 (1993).

⁴⁸ Daniel Susser, Beate Roessler, & Helen Nissenbaum, *Online Manipulation: Hidden Influences in a Digital World*, 4 GEO. L. TECH. REV. 1, 14–16 (2019).

⁴⁹ See *id.* at 38–41.

in the marketplace-of-ideas story is an optimistic prediction about what will happen when open information systems (“more information”) and authoritarian information systems (“censorship”) collide: As Darwinian conflict kicks in, truth (and by extension democratic self-determination) will prevail over falsehood (and by extension autocracy).⁵⁰ As Henry Farrell and Bruce Schneier show, that is too simple. Authoritarian information systems have developed sophisticated information strategies that leverage platform-based environments to undermine common knowledge about how democratic institutions function and, by extension, to destabilize the behavioral norms that lend such institutions continuing legitimacy.⁵¹ Such attacks, which are now well-documented, exploit platform capabilities for microtargeting, automaticity, and cascading, socially-networked information spread to stoke conspiracy theories and foster distrust—of government, of the “mainstream media,” of scientific consensus around topics such as climate change and the efficacy of vaccines, and so on.⁵² Powerful domestic factions that should have mobilized to defend these assaults on our foundational institutions instead have adopted weaponization techniques to further their own ends.⁵³ As such strategies become more powerful, they produce and amplify modes of public discourse about institutional actors that are incompatible with the knowledge structure of a stable democracy.

⁵⁰ Agreement with this prediction is deeply embedded all along the political spectrum. See Anupam Chander, *Googling Freedom*, 99 CALIF. L. REV. 1, 29–32 (2011); Matt Nese, *The Enemies of Free Speech Are Targeting the Internet*, THE HERITAGE FOUND. (June 10, 2019), <https://www.heritage.org/insider/spring-2019-insider/the-enemies-free-speech-are-targeting-the-internet> [<https://perma.cc/P3U2-UJAR>]; cf. Robert McMahon & Isabella Bennett, *U.S. Internet Providers and the ‘Great Firewall of China,’* COUNCIL ON FOREIGN RELATIONS (Feb. 23, 2011), <https://www.cfr.org/backgrounder/us-internet-providers-and-great-firewall-china> [<https://perma.cc/4D6J-HE9A>].

⁵¹ Henry Farrell & Bruce Schneier, *Common-Knowledge Attacks on Democracy*, Research Publication No. 2018-7, BERKMAN KLEIN CENTER FOR INTERNET & SOCIETY (Oct 2018), <https://pdfs.semanticscholar.org/4b52/376ddf73591114d597f992acdf108a1607a.pdf> [<https://perma.cc/MMF3-XVAH>].

⁵² See S. SELECT COMM. INTELLIGENCE, 116TH CONG., REP. ON RUSSIAN ACTIVE MEASURES, CAMPAIGNS, AND INTERFERENCE IN THE 2016 U.S. ELECTION (2019); Bradshaw & Howard, *supra* note 20; Hindman & Barash, *supra* note 20; see also MARGARET E. ROBERTS, CENSORED: DISTRACTION AND DIVERSION INSIDE CHINA’S GREAT FIREWALL (2018).

⁵³ See Paul Farhi, *Billionaire Koch Brothers Use Web to Take on Media Reports they Dispute*, WASH. POST (July 14, 2013); Suhauna Hussain and Jeff Bercovici, *Twitter is Suspending 70 Pro-Bloomberg Accounts, Citing ‘Platform Manipulation,’* L.A. TIMES (Feb. 21, 2020); Jonathan Swan, *Billionaire Father and Daughter Linked to Trump Shake-Up*, THE HILL (Aug. 17, 2016, 4:37 PM), <https://thehill.com/homenews/campaign/291772-billionaire-father-and-daughter-linked-to-trump-shakeup> [<https://perma.cc/X7UQ-3JQC>]; Julia Carrie Wong, *One Year Inside Trump’s Monumental Facebook Campaign*, THE GUARDIAN (Jan. 29, 2020), <https://www.theguardian.com/us-news/2020/jan/28/donald-trump-facebook-ad-campaign-2020-election> [<https://perma.cc/BT6G-BXZU>].

Each of these interests is compelling enough in its own right to warrant some degree of regulatory oversight. Their cumulative weight is considerably greater. But we have now arrived squarely at the problem of tailoring. Are there regulatory avenues that would safeguard the interests I have identified without doing violence to others that are equally important? Drafting such legislation is beyond the scope of this essay. Drawing on the analysis in Part I, however, I want to suggest two general sets of guidelines.

First, proposed legislation that is touted as targeting the dysfunctions of the platform-based, massively intermediated environment should stand or fall based on whether or not it actually does so—whether it responds to the failure modes of the platform rather than to abuses of distribution bottlenecks, to the types of remediable information complexity that frustrate the rational listener, or to the transparent venality of the intentional facilitator. Put differently, we should not expect interventions directed only toward the *largest* platforms, or only toward enabling individual choice about targeting, or only toward expanding DMCA-style liability or liability based on “reasonable efforts” at *post hoc* content removal, to accomplish much.⁵⁴ For similar reasons, platform initiatives for self-governance via “content moderation” should be understood for what they are: shiny, expensive distractions designed to stem the rising tide of criticism without undercutting the core platform business model, which depends on the relative profitability of immoderation. Oversight boards, internal appeal processes, and the like appeal to the lawyerly taste for process, but their significance is more performative than real.⁵⁵

By contrast, I have identified three structural features of platform-based intermediation that threaten the anti-factionalism, anti-manipulation, and anti-authoritarianism interests: predictive profiling and microtargeting based on behavioral and psychographic data; interface design to elicit automatic, precognitive responses; and algorithmic optimization to amplify patterns of cascading, socially-networked spread. Each of these features is amenable to systemic oversight, audit, and intervention, and platforms’ own actions confirm this. As platforms doggedly pursue ever more intrusive forms of behavioral and psychographic profiling, refine their interfaces to enable ever more seamless collection of user feedback, and continually tweak their algorithms to optimize both viewer engagement and networked information

⁵⁴ See, e.g., Mind Your Own Business Act of 2019, S. 2637, 116th Cong. (2019); Voter Privacy Act of 2019, S. 2398, 116th Cong. (2019); Deceptive Experiences To Online Users Reduction Act, S. 1084, 116th Cong. (2019); Biased Algorithm Deterrence Act of 2019, H.R. 492, 116th Cong. (2019); Honest Ads Act, S. 1989, 115th Cong. (2017); Danielle Citron & Benjamin Wittes, *The Internet Will Not Break: Denying Bad Samaritans § 230 Immunity*, 86 *FORDHAM L. REV.* 401, 415–23 (2017).

⁵⁵ For additional development of this point, see COHEN, *supra* note 14, at 135–36, 249–50.

spread, they also give the lie to the oft-repeated canard that their actions are intrinsically ungovernable.⁵⁶ It may well be “impossible to do content moderation well,”⁵⁷ but it is not impossible to imagine regulation targeted to those very different and more systemic failure modes, nor should it be beyond the pale of civil conversation among twenty-first century civil libertarians to do so.

A second set of relevant guidelines concerns the relative importance of different kinds of tailoring errors. All current versions of First Amendment scrutiny presume that the costs of mistaken instances of suppression (far) outweigh those of mistaken failures to suppress. That preference in turn rests on important assumptions about the nature and operation of the information environment—most notably, that injecting more speech into the marketplace is costly and that instances of low-value speech are readily ascertainable either by the rational listener or via intermediaries whose claims to authority the rational listener can readily assess. Those are not the properties of platform-based, massively intermediated information environments, and so the underlying presumption about error costs may warrant revisiting. As Frederick Schauer has explained, the First Amendment’s costs have always been

⁵⁶ See, e.g., Kirsten Grind et al., *How Google Interferes With Its Search Algorithms and Changes Your Results*, WALL ST. J. (Nov. 15, 2019, 8:15 AM), <https://www.wsj.com/articles/how-google-interferes-with-its-search-algorithms-and-changes-your-results-11573823753> [<https://perma.cc/FAA6-WDD2>]; Dana Mattioli, *Amazon Changed Search Algorithm in Ways That Boost Its Own Products*, WALL ST. J. (Sept. 16, 2019, 10:49 AM), <https://www.wsj.com/articles/amazon-changed-search-algorithm-in-ways-that-boost-its-own-products-11568645345> [<https://perma.cc/PY6V-9CQL>]; Tom Simonite, *Google Says It Wants Rules for the Use of AI—Kinda, Sorta*, WIRED (Feb. 2, 2019, 7:00 AM), <https://www.wired.com/story/google-says-wants-rules-ai-kinda-sorta/> [<https://perma.cc/V439-XYP5>]; see generally David Lehr and Paul Ohm, *Playing With the Data: What Legal Scholars Should Learn About Machine Learning*, 51 U.C. DAVIS L. REV. 653 (2017).

⁵⁷ See, e.g., Masnick, *supra* note **Error! Bookmark not defined.**; see also Mike Masnick, *Masnick’s Impossibility Theorem: Content Moderation at Scale Is Impossible to Do Well*, TECHDIRT (Nov. 20, 2019, 9:31 AM), <https://www.techdirt.com/articles/20191111/23032743367/masnick-s-impossibility-theorem-content-moderation-scale-is-impossible-to-do-well.shtml> [<https://perma.cc/XRN8-E6M2>]; Mike Masnick, *The Impossibility of Content Moderation Extends to the People Tasked with Doing Content Moderation*, TECHDIRT (Feb. 28, 2019, 9:43 AM), <https://www.techdirt.com/articles/20190226/18011641685/impossibility-content-moderation-extends-to-people-tasked-with-doing-content-moderation.shtml> [<https://perma.cc/WJE9-FUZY>]; Jillian York and Corynne McSherry, *Content Moderation is Broken. Let Us Count the Ways*, TECHDIRT (May 2, 2019, 9:31 AM), <https://www.techdirt.com/articles/20190429/15332242111/content-moderation-is-broken-let-us-count-ways.shtml> [<https://perma.cc/HJV9-Q25T>].

distributed unevenly.⁵⁸ As long as those costs did not threaten the overall stability of a system of democratic government accountable to the people and to the rule of law, however, they could be written off as the sort of collateral damage inevitable in a constitutional system designed to privilege liberty over equality and anti-subordination. Now that overall stability is on the table, however, it may be worth asking new and more probing questions about harms and costs.

Policymakers wanting to engage in a sensible discussion about tailoring and error costs, however, should remember that the project at hand entails designing effective oversight of behavioral conditioning and algorithmic amplification, and that the most effective forms of oversight will not consist of cumbersome, user-driven mechanisms for *post hoc* content removal. So, for example, we might begin by asking whether and under what circumstances we should agree to trade reduced scope for the viral spread of grass-roots political dissent against reduced scope for the viral spread of messaging about the need for armed insurrection in response to purported racial “replacement” or purportedly “rigged” elections. Properly understood, though, those questions are not about whether to jettison long-established principles designed to preserve breathing room for dissent. Rather, they concern the scope that a democratic system of government wishing to remain democratic should allow for microtargeting, manipulation, and amplification. The questions are important enough to warrant more than the usual knee-jerk responses, and they too should not be beyond the pale of civil conversation among twenty-first century civil libertarians.

IV. CONCLUSION

To appropriate a turn of phrase, the First Amendment is not a suicide pact. The mandate to preserve space for dissent, disagreement, and challenges to political and cultural consensus is vital, full stop. But the free speech imperative should not be interpreted to shelter the deliberate construction and fine-tuning of an information environment optimized to unravel the most basic preconditions for democratic self-government. It is platform functions and dysfunctions—rather than hierarchies and bottleneck effects, remediable failures of listener autonomy, or intermediary intentionality—that explain current threats to the anti-factionalism, anti-manipulation, and anti-authoritarianism interests. Platform functions and dysfunctions therefore should supply the frame for assessing constitutionally-required goodness of fit, and legislation appropriately tailored to the platform-based environment

⁵⁸ See generally Frederick Schauer, *Harm(s) and the First Amendment*, 2011 SUP. CT. REV. 81, 108–10 (2011).

and its particular democratic failure modes should be correspondingly more likely to survive review.