

# 9

## DATA-DRIVEN CAMPAIGNING

*Jessica Baldwin-Philippi*

Although the 2008 Obama campaign has been described as groundbreaking for its use of social media in academic and journalistic literature alike, its true breakthrough in campaigning was in the area of analytics and data strategy. This area has not exactly been overlooked, as the strategic choices involving data-driven campaigning have been detailed in popular press books (Issenberg, 2012) and how-to manuals (Delany, 2011, 2013) and have become central components of training young activists at major workshops like RootsCamp and conferences like Netroots Nation and RightOnline. The analytics teams that supported a variety of the Obama campaigns' work in both 2008 and 2012 have been mythologized by reporters as a cave full of geeks, hipsters, and hacks, insulated from the traditional campaign world of horserace coverage and process stories, and from which winning strategies emerged in computerized fluidity (Madrigal, 2013). While these practices mark important changes to the work of political communication and campaign strategy more broadly, the tendency of both journalists and academics to focus on the winners and innovators in order to discern how political communications strategies change and what tactics are most successful does not tell the whole story.

This chapter aims to describe the state of data-driven campaigning as it has been taken up by campaigns that have fewer resources—in skills, labor, money, and sometimes all three. Following an overview of what data strategy has looked like over time and the concerns that have accompanied its rise, this chapter gives an account of the evolution and status of data campaigning across Field Operations—the team behind voter registration and often volunteer outreach, making contact with likely supporters via phone and in-person canvassing, and getting out the vote (GOTV) on Election Day—and Communications—those who craft email, traditional and digital advertising, and most often have a hand in producing social

media content. Overall, it argues that data-driven campaigning is a more limited practice than many popular accounts of contemporary campaign strategy acknowledge, and that beyond presidential campaigns and the most competitive of senate runs, staffers limit data-driven strategy to field operations and advertising.

As part of a volume about the role of data and analytics in political communication research, this project takes the somewhat counter-intuitive approach of using qualitative methods to investigate the topic. In doing so, I take an epistemological approach to data, investigating how campaigns think about data, the conditions under which data produces meaningful knowledge, and the types of knowledge it produces, rather than looking at data itself as the object of study. Moreover, I focus on the everyday ways that data is (and is not) used, discussed, requested, or ignored by campaigns. This chapter is not merely about the contemporary practices of campaigns, but how and why those practices have emerged. In doing so, I find two somewhat divergent stories. On one hand, presidential and advanced senate-level campaigns are staffed by vocal advocates of data-driven practices and the most analytically advanced activists. On the other hand, more local campaigns—congressional, gubernatorial, and even some of the less competitive senate campaigns—are not engaging in these practices in all facets of campaigning. This chapter goes beyond documenting these differences, and highlights where and why such divergence has emerged. Far from a simple story of merely needing data skills, this chapter uncovers other variables that prevent the uptake of data-driven analytics in certain types of campaigns and particular aspects of campaign strategy. In doing so, it also speaks to lessons that the field of political communication can glean from attention to this gap between the data haves and have-nots.

## The Development of Data-Driven Campaigning

The roots of data-driven campaigning go much deeper than digital campaigning. Beginning with the rise of public opinion polls in the late 1940s and early 1950s (Berelson, Lazarsfeld, & McPhee, 1954; Lazarsfeld, Berelson, & Gaudet, 1948), data has been used to inform political campaigns of what issues are important to voters, which stances are most popular, and which arguments are persuasive. Over time, this data also provided insight into how different populations viewed issues and could be persuaded and/or mobilized to vote on Election Day. In the 1980s, campaigns began using data from Census Block Groups to combine economic data with racial demographics, education levels, and religious beliefs to construct archetypes of Americans, which were then understood to have political valences. For instance, “Furs & Stationwagons” living in “expensive new neighborhoods in the greenbelt suburbs of nation’s major metropolitan areas,” their empty-nester equivalents, “Pools and Patios,” or the recently-divorced “New Beginnings” set (Klein, 1989). During the 1996 election, the Clinton campaign further refined such tactics by collecting highly segmented opinion polls and lifestyle data to

construct archetypes of voters with varying levels of persuadability and issue focuses. “Soccer Moms” were suburban women who the Clinton campaign saw as likely persuadable swing voters, while a reconstructed version of “Pools and Patios” voters were more likely to vote Republican.

Contemporary visions of data-driven campaigning combine these tactics of population defining and targeting with improved database technology that can provide more data to parse, more precisely slice and re-combine populations, and even test how messages perform across groups. As early as 1992, the California Democratic party began organizing voting records, matching that with other available information, and adding its own data about voter contacts (phone calls and canvassing). In the 2004 election, the Bush campaign engaged in extensive efforts to identify potential and likely voters, and built complex models to test and improve turnout (Kreiss, 2016). Over a decade later, these digital platforms have made the collection and organization of information relevant to political decision-making and mobilization of all types easier. While most mobilization efforts are still aimed at turning out voters on Election Day, other, higher-dedication actions such as getting individuals to donate or come into a campaign office to volunteer—and moving them up the so-called “ladder of engagement” to increasingly important and time-consuming actions—are also part of mobilization efforts and are increasingly data driven. Cookies record which parts of campaign websites we click on and when, email systems can organize individuals into separate groups and document which individuals opened which emails and if links inside of the email were clicked, and social media platforms construct our identities out of the information we willingly provide in profiles, the articles we click on, and the preferences of individuals we connect with, and then target political advertisements accordingly. Even more advanced campaigns go beyond relying on targeting to perceived political values, and empirically test particular messages, images, and calls to action with small samples of users before sending the perfectly tailored versions of each message to all available parties (Baldwin-Philippi, 2016). Embedded in this drive toward increasingly precise targeting is what Beth Coleman (2016) has called the “hegemony of big data,” or the belief that more data and more specific data will allow campaigns to truly know their audience, and therefore know how to persuade or mobilize them. Although Eitan Hersh’s (2015) book *Hacking the Electorate* has demonstrated that the data that is most valuable to campaigns is that which is publicly available and much more overtly tied to past political behaviors and broad demographic information—voter registration, census data, vote history, rather than consumer or lifestyle data—microtargeting and subsequent message testing makes up a significant portion of the discussion of campaign practice.

Even prior to the precision offered by data and analytics that contemporary digital campaigning has enabled, concerns about the democratic impacts of targeting messages to specific populations have existed. First and foremost of these concerns is the fear that with the rise of microtargeting, we will miss out on information that we may disagree with, thereby leading to further polarization and a

diminished level of informed citizenry (Pariser, 2011; Sunstein, 2007). The production of a so-called filter bubble in which people find themselves immersed in content they are likely to agree with is not simply a function of gravitating toward information you agree with or avoiding information you do not, though digital information consumption practices certainly reflect this tendency to selectively expose oneself to agreeable content (Garrett, Carnahan, & Lynch, 2011; Stroud, 2011). It is also a function of the technological platforms campaigns and citizens alike rely on. Facebook allows individuals to mute other users and group accounts, as well as particular publications. Facebook, YouTube, Twitter, Instagram, and so on, also provide candidates with highly targeted options for ad buys, including categories that are based not only on political leanings, but of hobbies and interests as well. As campaigns not only target people who are likely to be persuaded or mobilized, they also avoid or segment those who are unlikely to support the candidate, thereby impoverishing certain populations of information. Both of these practices harm the potential citizens have to become informed, and therefore hinder the potential for democratic deliberation and debate. Data-driven practices like microtargeting are not just about ensuring the public is informed so that deliberative democracy can flourish, but also about ensuring opportunities for participatory democracy are present. Just as certain populations can be segmented and end up with less information, the same practices of segmentation leads campaigns to avoid mobilizing certain individuals and groups. Moreover, as Philip Howard (2006) has argued, these tools also reward the loudest voices, rather than the most informed, thereby enabling the rise of “thin” citizens, who “can respond quickly to political urges and need not spend significant amounts of time contemplating political matters” (p. 185). That said, practitioners often argue that voters are more likely to respond to and become mobilized by content they are interested in—especially in an age of information overload—thereby giving microtargeting a valuable role in constructing a more participatory democracy (Issenberg, 2012).

As described, data-driven campaigning strategies are available to campaigns, but central questions remain: Where is data-driven campaigning occurring, and where is it faltering? The following analysis describes how Field and Communications operations within campaigns have made use of data-driven strategy, and where such practices remain lacking. Rather than treat lack of adoption as a failure, this chapter highlights the reasons behind some campaigns’ lack of data-driven strategy, and in certain aspects of campaign communications in particular. This chapter then discusses the implications of these findings on the field of political communication, asking what this means for scholars who study campaigns, and scholars who study emergent and changing strategies. Rather than using big data or analytics to conduct my inquiry, I inquire about what these things mean for campaigns.

To answer these questions, this research does not look to big data itself, but to a qualitative investigation of the practices, strategies, and tactics of campaign practitioners and their reflections on these practices, as well as content from training materials and campaign manuals. As a result, it draws on data spanning

from 2010–2014, including in-depth interviews of over 40 campaign consultants following the 2010 and 2012 campaigns, brief (~ 20–30 minute) interviews with an additional fifteen staffers and consultants in 2014—nine of whom were staffers in single campaigns in 2014 (three in senate races, six in congressional, across six different states) and six of whom were consultants working on multiple campaigns across a variety of states, often including lower state-level candidates as well. It also uses content presented at professional consulting conferences and written manuals by leaders in the field over the course of those four years. Additionally, as it draws on shifts in analytics-based strategy and the rise of the culture of testing over time, it draws on prior fieldwork as a participant observer inside an Illinois race for federal office in 2010 (Baldwin-Philippi, 2015).

### Field Operations: Data-Driven Voter Contact

Historically speaking, Field Operations have the greatest experience with data-driven politicking. While public opinion polls may contribute to how campaigns construct persuasive messages, the use of publicly available, privately purchased, and privately collected data is where the most meaningful and widespread use of data exists. Using combinations of the above data types, campaigns have long used such information in their “ground game”—to predict which individuals and households campaigns need to directly contact in order to ensure they turn out on voting day. While stories of declining ground efforts and the rise of mass media campaigning define the 1980s and 1990s, the late 1990s marked the beginning of increased interest in expanding turnout efforts. Campaigns began devoting significant amounts of money and labor to GOTV efforts in the early 2000s, which were themselves increasingly based on rigorously-tested experiments concerning what worked when trying to get people to the polls (Nielsen, 2012). Early academic experiments conducted during the 1998 midterm elections tested how effective various forms of voter contact were, pitting direct mail, phone calls, and in-person canvassing against each other, and were released just before the 2000 election, to much attention from the consulting class (Gerber & Green, 2000). Daniel Kreiss (2016) goes into great detail discussing the histories of database technologies and use of data on the right and left, explaining how, like the California State Democratic Party began to do in 1992, the Republican National Committee (RNC) had a relatively robust system of sharing voter records and other data that was then used to canvas and target individuals for GOTV efforts. As his book, *Prototype Politics: Technology-Intensive Campaigning and the Data of Democracy*, details, this system, called Voter Vault, was developed in a federalized manner in 1995, and by the 2002 midterms was turned into a major database where state parties could connect directly with an interface that would allow them to work with (and add to) data from their local voter files. This system widely outpaced that of the Democratic National Committee (DNC) during that time period, and marked a major development in data at the national level. With this buy-in from parties,

databases began to be widely shared, accessed, and added to in a decentralized manner. Following subsequent failures of the RNC to keep the data and technology up to date and the DNC's increased efforts in the area, the DNC caught up to and surpassed the RNC.

In 2004, the AFL-CIO conducted experiments on which unions and which kind of members were more likely to be persuaded by political appeals and ads from leading Republicans (Issenberg, 2010). As the resulting empirically proven tactics were taken up by campaigns on both side of the aisle, and data-driven GOTV tactics and targeting took off in earnest, the political left decided it would be in their interest for more of these findings to be proprietary. In 2007, the Analyst Institute was established to do just that, and brought academic rigor to the experimental field-testing of voter-contact methods. Over the next seven years, the Analyst Institute continued to test methods of voter contact (text messages are helpful (Malhotra, Michelson, Rogers, & Valenzuela, 2011)), and actual scripts to increase turnout (social pressure works (Gerber & Green, n.d.)). Organizationally speaking, while Obama's digital and analytics staffers contributed to many projects—using social media data to target emails and other communications, fundraising, and so on—their crowning achievements in both 2008 and 2012 were turnout models and data-driven organizing. Although it was infamous for its failure, the 2012 Romney campaign's ORCA project—a data-driven Field project designed to show the campaign precisely who had turned out in real time on Election Day—became a cautionary tale against failing to properly invest in technological infrastructure.

Importantly, the use of such data for field operations has not been limited to presidential races—in fact, its power has laid in its usefulness for state-level and local races. While many campaign innovations, such as learning best practices for social media or email writing, take time to trickle down from the presidential level to that of congressional or gubernatorial races (Baldwin-Philippi, 2015), using data to personalize voter contacts was occurring in congressional level elections by 2008 in a way that was more than haphazard, long before they adopted social media (Nielsen, 2012). Additionally, on a more ad-hoc basis, campaigns have long kept lists of “1s,” who they know to be likely to attend an event or lend a last-minute hand if needed, as well as “2s” and “3s” who are reliable bets for some volunteer or donor activities. While campaigns at the congressional level and all but the most well-supported senate races lack the type of data that would allow them to make the predictive analysis regarding what type of persuasive messages were likely to work (which the Obama 2012 campaign did do), major party candidates have had access to voter data that is decently well-organized and pretty complete since 2008. Importantly, public records, more than specific lifestyle data, are what have significant impact on strategy at all levels of campaigning—including the presidential (Hersh, 2015). At the congressional level, as early as 2008, campaigns were not only using vote history data to decide who to directly contact by phone, in-person canvassing, and/or mail, but were adding to the database of information

about those voters by recording how likely they were to vote for the candidate or entire party ticket, and whether or not they would be willing to volunteer or donate. In 2010, much of the data added to the system was first recorded by volunteers in analog, and was subsequently entered manually. With the proliferation of smartphones and further decentralization of this data, volunteers can add to the file at their fingertips, and the practice is now nearly universal. Daniel Kreiss's (2012, 2016) work explains the differences in database infrastructure and the divergence in the effectiveness of data-driven strategy across the right and left in great detail, but what is notable is that data is used across the board, even if the left has been more adept in its uses over the past decade.

### **Communications Operations: Email, Social Advertising, and Social Media Content**

In a modern congressional campaign, the communications team sends out multiple messages directly to potential voters and supporters every day. From national ad campaigns to responses to Facebook comments, mass emails to tweets, campaigns communicate constantly. Increasingly, we have heard about campaigns' ability to target such messages. The Obama 2012 campaign used new, increasingly personal data from the television ratings industry to develop data-driven practices that allowed the campaign to target television ads by personal political information rather than the traditional television industry terms—likelihood of voting, level of information they'd received, and political leanings, rather than the demographics of viewers of certain shows or genres (Davidsen, 2015). As companies like Facebook and Twitter depend on ad revenue, they have made possible targeting very particular groups of people based on interests that are both self-reported and determined by content sought on the platform, friend networks and interests, and traditional demographic information like location and age. Most campaigns use customer management systems (CRMs) that can organize email lists and facilitate targeting and testing of messages. These practices all fall under the banner of what many consultants have, since 2010, called a "culture of analytics" or a "culture of testing" that advocates for rigorously testing messages across many platforms in order to ensure they are not only well-crafted, but well targeted, and also provide campaigns with more information about what works moving forward (Baldwin-Philippi, 2016).

#### ***Mixed Inbox: Targeting without Testing***

With little effort or infrastructure, campaigns can purchase subscriptions to off-the-shelf Customer Relationship Management (CRM) systems that allow them to organize recipients and target messages based off of information they ask citizens to provide, such as political interests, types of events they might be interested in, and basic demographic information. CRM systems can also collect data on whether

an email was opened, if it was shared in other social media spaces, and/or what links were clicked. Such systems return the data concerning which messages were opened and otherwise responded to in varying degrees of detail, often allowing campaigns to easily see which emails were opened, or which populations did the most opening. Although campaigns can learn a lot from eyeballing the popularity of data like open and click-through rates, consultants and staffers with more advanced analytics experience have begun to advocate for what has been called a culture of testing, which involves scientifically testing messages, rather than merely eyeballing trends. Major proponents of the culture of testing also use consulting conferences to advocate for more advanced testing involving randomized controlled studies to train activists and consultants to be efficient in how and when they communicate. Although the language of cultivating a culture of testing is seemingly more widespread than the actual use of such tactics (Baldwin-Philippi, 2016), advocates argue that their use allows campaigns to improve their outreach by small, but meaningful numbers, which is what will result in more persuasion, dollars raised, and ultimately more electoral victories (Michelson & Nickerson, 2011; Schwartz, 2009). Many of these tactics are based on the field experiments pioneered by Field teams to mobilize voters to turn out on Election Day, and have since been applied to mobilization efforts surrounding getting supporters to donate, volunteer, or take online action such as petition signing.

While presidential campaigns have made use of these practices since 2008, congressional and even senate level races' use of rigorous scientific testing is limited, despite their professed enthusiasm for analytics. Generally, communications directors for Senate and Congressional campaigns were well versed in targeting email lists according to issue interests and geography. Campaigns will send invitations to events to those they know are local, based on self-reported zip codes while sending fundraising requests outside of their district, and campaigns will send out updates about an environmental policy or event with environmental groups to those they who have checked a box stating they were interested in that particular issue. Well-funded congressional campaigns were practicing this regularly as early as 2010, even though they also sent out mass emails to their entire lists as well for things like fundraising, calls to action, and occasional campaign updates such as new ads and endorsements (Fieldnotes Aug 25, 2010). The practice of sending some emails to an entire list persisted at the congressional level, where they were largely used to update supporters on events like fundraising dinners with major elected officials, and "newsletter-type" emails with multiple updates (Personal communication Dec 15, 2014). Interestingly, even at the presidential level, a common critique of the 2016 Trump campaign was that the communications office was not even targeting its email list at this basic level, and engaging in practices that made it likely that messages would go to spam (Kaye, 2016; Samuelsohn & Vogel, 2016).

Calls to go beyond this simple list splicing and move toward testing messages are relatively widely discussed in professional trade publications like *Campaigns & Elections* and major conferences such as *Netroots Nation*, *Conservative Political*

*Action Conference (CPAC)*, and the now-defunct *Right Online*, as well as by leading consultants. Outside of the presidential level, those who do attend to controlled sample testing are, overwhelmingly, either consultants or staffers for the most competitive of senate races. With a recent organizational study showing that roughly 20 veterans of the Obama campaign who focused on digital and analytics went to down-ballot races in 2014 and 2016 (Kreiss & Jasinski, 2016), it makes sense that these spaces are where the most innovative tactics would take place, leaving the rest of campaigns to less rigorous methods of message testing.

In most campaigns, these practices are simply not a predominant aspect of local races. In 2010, only the most advanced campaigns at the senate level were testing those messages in order to determine what type of subject header, content, and images were best at gaining attention or mobilizing constituencies (Baldwin-Philippi, 2016). By 2014, those at the congressional level and less competitive senate races had made little headway in this area. At the congressional and senate level in 2014, few consultants and staffers engaged in randomized, controlled sample testing in 2014, though some did regularly compare two or three versions of a message and send the one that worked best to the whole list. This practice, known as A/B testing, does not separate test subjects into controlled samples or isolate particular variables of the message. Campaign consultants in particular, rather than staffers on single campaigns, reported A/B testing, and most used it for both website and email content—five out of six I spoke with used this tactic for any type of message, from website content to social media ads, although only three reported using it for email. Although staffers were far less likely to use A/B testing, it was still a viable tactic, with three of the nine staffers interviewed engaging with such comparative analysis. Two of these worked at the senate level, and each of them reported that they mostly A/B tested their website content, rather than emails. Despite using this tactic, staffers (especially at the congressional level) were unlikely to describe conducting such tests in a routine or ongoing way.

### ***Advanced-level Advertising***

Advertising tells a slightly different story, as campaigns' facility with targeting ads far surpasses any other communications efforts, at all levels of election. While social media were not always ad-supported, the presence of advertising in the digital campaigning landscape is unavoidable. Facebook has been home to a robust ad system since 2008, and Twitter has since April of 2010.<sup>1</sup> Because social media platforms subsist on advertising revenue, they have long lured businesses with the very fact that they have enormous amounts of data about their users' interests and behaviors and can target them accordingly. Over time, platforms like Facebook and Twitter have made ads look more and more like the content produced by other users, changing their placement in the user interface and using small, less noticeable disclaimers about promoted content. Broadly, campaigns have more experience with advertising within Facebook than Twitter, due to Facebook's

audience size and the perception that Twitter is best for talking to journalists (Hamby, 2013; Kreiss, 2012) and driving the campaign's narrative.

As early as 2010, even congressional campaigns were much more familiar with targeting ads in Facebook and looking at the ad analytics Facebook provides in order to assess whether an online ad was succeeding, and making the appropriate changes from these findings. For instance, campaign staffers who did not A/B test any email or web content were familiar with targeting Facebook ads by location, age, education, gender, and other demographic features. In fact, they used the advertising platform easily as early as the 2010 midterm election (Fieldnotes, August 25, 2010). Moreover, Facebook allowed microtargeting based on keywords that users had entered as interests of theirs, and specific places users work. Opportunities for microtargeting have only gotten more specific as Facebook moved past using only user-provided information, and began to assemble their own "shadow" profiles of users based on what we click on, what we like, and who our friends are. As one congressional staffer explained, "Facebook has always made it so easy, with drop down menus of demographics and interests and everything. Now there are even more options [than in prior years]—you can boost posts, not just ads even" (Personal communication, December 18, 2014). As early as 2010, another widely used method of targeting that Facebook provided was to segment populations of those who were fans of a page, those who are not, and those who have "friends" who are fans of a page, enabling campaigns to target new potential supporters while also adding social pressure by showing them which of their friends have already liked their campaign. Similarly, as early as 2008, campaigns began to match their voter files to ad networks, Yahoo!, and Google, so they could better-target ads on search engines, and on scores of websites that subscribe to various ad networks. The ability to microtarget has only gotten more precise over time. By 2015, campaigns could even target individual "political influencers" across the political spectrum through Facebook's ad interface, and the Trump 2016 campaign's digital team claimed that highly targeted Facebook ads were key to their success (Green & Issenberg, 2016).

In all of these cases, the corporations selling the ad buys have made it cheap and, even more importantly, easy for campaigns and have shared best practices and advice for how best to use their platforms for targeted advertising. One staffer from the congressional level explained that corporations not only made their ad interfaces intuitive, but have doubled down on offering support to encourage campaigns to advertise within their platform: "There have always been newsletters or listservs I read for best practices of all kinds of messaging, but Facebook ads were always pretty easy. And now Facebook even emails advice and best practices if you sign up for their list" (Personal communication, Dec 3, 2014). Campaigns could also use the data they receive from their own fans to design strategies that expand beyond advertising. In 2012, the amount of data that Facebook provided to campaigns about anyone who had liked their page allowed the Obama campaign to target emails encouraging people to vote by telling them which of their friends had pledged to do so (Judd, 2011).

While the 2008 Obama campaign used Facebook to both broadcast and target their messages via public posts and advertisements respectively, the platforms and strategies for doing so have changed dramatically over the past eight years. In 2010, the line between public and targeted content was clearly drawn—posts were public (unless sent directly and privately to individuals), and advertisements were available to be targeted based on users' location, age, gender, and interests. Though many campaigns engaged in both tactics, they were sometimes surprised and often frustrated at the difference in affordances, and their inability to use the targeting tools they used for advertising to target regular public posts or messages (Fieldnotes, Aug. 31, 2010; Sept. 1, 2010).

### ***Social Media Analytics: Using What's There***

While figuring out if ads worked, and how competing versions tested against each other was relatively easy for campaigns, understanding what worked for content of social media posts has been more difficult for campaigns. Of course, testing particular social media messages ahead of time is not feasible—the posts are public and dispersed on a mass scale. Instead, campaigns use social media post-fact to determine if messages worked, although defining what success means in this space has been difficult. Social media platforms like Facebook and Twitter are also continuously developing analytics and dashboards to help their users—especially those related to consumer marketing, who also buy ads—make sense of these numbers. They are also developing improved ways to help campaigns expand the reach of messages and even target them, including both paid and free options.

If campaigns like looking at click-through rates of emails, they love looking at the social media metrics provided in the platforms' dashboards—for instance, Facebook Insights' clear illustration of which posts got the most hits, as represented as overall bar graphs, and by impressions per day. One extolled their virtues of being immediately available and easy to decipher, exclaiming “Of course! They're right there! You can just take a quick look and see what type of messages perform well” (Personal communication, December 18-b, 2014) when asked if she looked at them. Another laughed when asked the question, noting “I was a little compulsive with them. You want people to see your messages, and to be able to see that feels good—as opposed to like calling, where who knows what the guy who hangs up on you will do or if he'll vote.” (Personal communication, December 13, 2014). Even though likes and retweets can tell staffers little about what Election Day turnout will be like, metrics provided some window into how potential voters would respond, and that was gratifying, even when anxiety-provoking. This staffer explained this feeling of anxiety as well, saying he “had to say ‘Ok, you can't look until tomorrow morning’” so he could both unwind when home, or work on other things that required more focus.

When looking at these analytics, it's important to acknowledge that even as campaign staffers and consultants say “they show what works,” the very definition of success is complicated and not measured particularly well by the available analytics.

At their core, these analytics focus on popularity, highlighting the amount of impressions content makes, how many times it is liked or favorited, and how often it is shared about other metrics. Visually, the native analytics of Facebook overwhelmingly emphasize these very measures. In each of these, the number of likes, the reach (total impressions), and “engagement” (a combination of likes, comments, shares, and how many times links are clicked) are provided. These engagement analytics begin to provide insight that is deeper than popularity of topics or issues, but also require more advanced strategies in combination with other forms of analytics to understand if other action such as donating, or signing up for a newsletter are taken. Facebook highlights notifications specifically for likes and shares and bar graphs that visually represent the amount of people reached. Although popularity and reach of messages are important to campaigns—not least of all to seem well supported—these messages are unlikely to persuade undecideds, or those leaning the other way. Instead, social media messaging—and digital messaging more generally—has largely been viewed as a mechanism to deepen commitment and mobilize those committed supporters (Baldwin-Philippi, 2015; Bimber, 2003; Vaccari, 2008). And social media platforms are particularly bad at measuring that. The analytics provided by platforms like Facebook and Twitter do measure instances of mobilization or commitment-building, such as engagement, or if audiences click on links provided, but the systems themselves cannot track analytics of more action-oriented behavior, such as which users donate once a link is clicked, or how long they stay on that page. While Twitter provides data on a user’s audience through a “followers” section of their analytics and a very skilled campaign could perform social network analysis of its users, that information harder to access than analytics displaying an account and individual tweet’s impressions, retweets, favorites, or an account’s mentions. While campaigns know that popularity is not necessarily the most important metric, it is one they often fall back on.

Just as the material constraints of privately owned platforms impacted campaigns’ ability to control and censor public commentary, they also constrain campaigns’ ability to work with analytics. The data that platforms like Facebook and Twitter automatically provide to users—especially the data that they choose to visually represent and make the most easily interpretable—will be the data that campaigns turn to, and that therefore drives content creation. Popular social media analytics packages such as SproutSocial, CrowdTangle, Attentive.ly, and so on, provide more complex metrics of “engagement” than those baked into social media platforms themselves, like Facebook Insights. In these third-party tools, algorithms rank combinations of measures such as likes, shares, attention, users’ networks, and so on, and sometimes compare them with other similar pages. These uses of analytics are certainly an important beginning step into testing messages and engaging analytics more rigorously, but their limited use results in attention to the type of content that encourages campaigns to focus on messages that are popular, rather than mobilizing.

What campaigns are doing with social media analytics is rather mundane, and not particularly helpful to creating more nuanced messages, but it is also in the

platforms' interest to develop tools that enable campaigns and consumer marketers to make better use of the vast quantities of data social media platforms collect. In one major change, by the time the 2012 campaigns occurred, Facebook had blurred the line between public posts and ads, giving campaigns the ability to "promote" (now "boost") individual posts to audiences that could be targeted based on user location, age, gender, and interests. Social platforms like Facebook and Twitter have continued to blur the line between advertising and content by allowing campaigns to pay to "promote" their posts to larger audiences and target the populations who make up that audience. While these posts are still publicly visible, and therefore subject to checks by journalists, supporters, and the opposing campaign alike, campaigns can now extend the reach of messages they choose, and do so to populations that are microtargeted. Still, some campaign staffers are skeptical of such options. One reported, "It's just like \$3, so of course I did it sometimes, but I never really felt great about it. Like, what do those extra people reached mean? I'm glad when it says more, but is that worth paying for regularly? Is this just Facebook making us think it works?" Interestingly, this tactic brought forth skepticism that was previously not discussed among consultants and staffers.

### Why Are Local Campaigns Falling Behind in Communications?

In many cases, new campaign strategies trickle down from the well-funded presidential campaigns that are staffed by well-known and successful staffers to more local races, staffed by more local politicians with less experience and less exposure to cutting-edge best practices. Of course, there are outliers, but this trickle-down narrative wherein innovation happens at the presidential level and adoption occurs first by well-funded, competitive campaigns, then by local campaigns has held true for how campaigns made sense of how to navigate the type of content to post in social media, how to craft websites, and so on (Baldwin-Philippi, 2015; Foot & Schneider, 2006; Nielsen, 2012; Serazio, 2015; Stromer-Galley, 2014). Most of the arguments for why local campaigns are falling behind in the move toward microtargeted and analytics driven emails and social media messaging assume that staffers and consultants at this level lack the skills or professional experience to believe in data-driven strategy.

At a certain level, this is true, but it does not tell the whole story. Daniel Kreiss's study of where presidential campaign staffers work off-election years shows that they do go to major senate races, but are very unlikely to work at the congressional level, using their skills and knowledge to improve the tactics in their races and leaving others to catch up over time. Races at the congressional level, therefore, do lack the individuals associated with data-driven campaigning, but they do not necessarily lack the skills to do so, or the more general cultural norms needed to want to invest in doing so. When speaking with staffers of and consultants for senate and congressional level races, nearly all of them report wholeheartedly supporting moves toward data-driven campaigning. Moreover, not all local races

are lacking skilled practitioners, as hundreds of local activists, advocacy workers, and campaign staffers have been trained by data-oriented grassroots organizations like the New Organizing Institute and Wellstone. What this means is that at the level of senate and even congressional campaigns, staffers often have these skills, but are unable to make use of them for other reasons. In particular, more local campaigns stress the amount of labor data-driven practices take, the limited payoff, and the availability of other information that gets “close enough” to data-driven practices.

Campaigns, especially local ones, are staffed by very few people, and the time of those who decide and implement strategy is stretched very thin. At the level of congressional campaigns, the best-funded campaigns generally have fewer than fifteen staffers, with many campaigns working with fewer. For senate races, the higher end of these employ around 25 paid staffers. While volunteer labor and party resources provide reinforcements, those individuals are not usually tasked with strategic positions or the power to craft much content. With the time of digital and communications directors stretched so thin, they are unlikely to have the time to devote to data-driven practices. One staffer who had been trained in and previously engaged in email targeting and testing for an advocacy group reported, “we just didn’t have the bandwidth—or the time really—to write a bunch of different emails [to test against one another] (Dec 15, 2014b).” Despite knowing that he could increase the number of people who read an email or click on its link, he was not willing to use his valuable time to write multiple versions of an email to test whether different images, narrative frames, or tone would change the results. Similarly, as another communications staffer for a congressional level race explained, “I don’t have to take time to set something up before posting a message [...] I can look at the numbers [of who liked or shared social messages] when I have the time” (Personal communication, December 3, 2014).

Moreover, the ultimate payoff of taking the time to do so is not immediately beneficial to campaigns at local, and even statewide levels. The success of data-driven campaigning in many ways relies upon large sample sizes. When the most successful of messages increases your desired outcome by 3 percentage points, and most of the time testing only increases desired behaviors at a rate below 1%, the payoff is markedly less for local races. When the Obama campaign nets a 1% increase in donations on a list of about 20 million, that number is massive; when a local race has a list numbering in the single thousands, 1% makes much less difference. The tradeoff between time and reward is fundamentally different for local races, and decisions to forego data-driven email and social media tactics can be rational, considered responses to the high cost and lesser reward they offer smaller campaigns. The constraint of Election Day looming as the ultimate arbiter of success and failure becomes especially visible when considering how much more advanced the uses of analytics-based messaging activist organizations—even relatively small organizations—engage in (Karpf, 2016). When an investment in labor

can pay off later on in future campaigns, months down the road, such practices become much more widespread.

Campaigns have also found workarounds for adopting what they assumed were best practices. One very popular method of assessing social media content, as discussed earlier, is using available analytics that measure popularity and reach, rather than more nuanced metrics that measure which messages are successfully being circulated, shared, or liked. As discussed earlier, campaigns of all sizes had staffers who paid attention to the analytics that social media platforms provide users. A staffer who said she didn't have the time to test messages noted that she could come to social media metrics after content had been posted, see if it became well-shared, well-liked, or even just seen by many people, and "still use that [finding] the next time we need to post a message" (Personal Communication, Jan 12, 2015). These metrics are how, as the 2010 midterm elections began to progress, campaigns realized that posting long, text-heavy, issue-focused content in social media did not get much attention in the space. As one digital consultant handling many races explained, this learning process happened over the course of the campaign, and by the end of 2010, they realized that "anytime we dabbled with actual policy, it was a disaster" (Personal Communication, March 4, 2011). While close attention to metrics or testing email messages might have revealed that sooner, campaigns caught on nonetheless.

Sometimes foregoing data all-together, campaigns often looked to each other for best practices, and copied the style, tone, and content of campaigns that are better resourced, assuming knew what they were doing. In 2010, congressional communications and digital directors constantly referenced Cory Booker's use of Twitter as something they wanted to emulate in order to reach out to constituents. "Look at [our opposing campaign's] pictures. You need to use images that are better framed and use filters, like [our opponent]" (Fieldnotes, Oct 18, 2010). Another staffer who worked on a senate race in 2014 explained how they had adopted certain email writing practices, saying "You know how Obama's emails [subject lines] are all short—like, "Hey..."? We know that works better. So we started doing short subject lines too" (Personal Communication, Jan 16, 2015). Other work investigating the use of social media strategy in the 2016 presidential primaries showed that even at this level, staffers from lesser-known candidates would look to the perceived leaders in the field (McGregor, Kreiss, & Lawrence, 2016). For instance, if a campaign of either party saw the Clinton campaign creating a new style of image or content, they would be imitated, because the perception was that their campaign knew best practices. In these cases, campaigns can decide to forego testing messages (whether it is outside their skill level or at too high a cost), while thinking of their strategy as benefitting from the move toward data-driven messaging.

Of course some opportunities to use new strategies and analytics are missed due to campaigns' inability or lack of vision about the importance of such methods. Even as campaign staffers and consultants extol the virtues of data-driven

practices, the culture of campaigns can privilege tried and true methods and gut instincts that have worked in the past, rather than pay attention to what new methods are seemingly working in the present. But these cultural clashes between traditional and new campaign strategies are not limited to the local level. In both 2008 and 2012, the Obama campaign's data and analytics department clashed with more traditional media and communications departments. Staffers from competitive senate races report repeated arguments between Communications and Digital operations. The Romney campaign famously required 22 people to sign off on anything the campaign tweeted, to the great frustration of the Digital team (Kreiss, 2014).

### *Implications for Studying Data-Driven Campaigning*

As the study of data-driven campaigning reaches beyond accounts of what the most successful and well-funded of presidential campaigns are doing, we not only learn how strategies differ across levels of campaigning, but how and why they spread—or do not. At a fairly surface level, data like that presented here tells us that the whole terrain of political campaign strategy does not look like that of presidential campaigning. While presidential races tell us much about cutting-edge tactics, it is the midterm elections where these practices get picked up or left behind, and they play a vital role in continuing, evolving, and sometimes changing or ignoring strategies from presidential campaigns. These campaigns are home to many who continue to work in local politics and advocacy organizations, and we miss a significant amount of what political strategy looks at by looking only to presidential campaigns.

Moreover, the assumption is that innovative, productive tactics and strategies that are developed at the presidential level will trickle down as skills or tools disperse. Given that testing messages and using analytics does rely on a particular set of skills, it is sensible that we might attribute more local campaigns' failure to adopt a product of missing skills or because those in charge of campaigns are rooted in prior norms. Consequently, that tactics that work at the presidential level will, over time, work their way into down-ballot races. This research points to flaws in that thinking. The trickle-down model depends on the only variable precluding adoption being knowledge of the strategy or skills/tools needed to enact it.

This work points to alternative reasons, such as differing constraints on labor and resources, that makes congressional and all but the most competitive senate races unlikely to make such choices. We discuss email as a mode of communication that is free of cost, and that does not tell the whole story. While email is, obviously, cheaper than advertising, it does take significant labor to produce good, effective emails, and costs the communications director and digital director valuable time—a resource they cannot fundraise to get more of. Gubernatorial, congressional, and senate races (to say nothing of mayoral or state senate) are unlikely

to raise enough money to simply hire more staffers to make up for this labor, as presidential campaigns could do. Moreover, while the cost is higher to down-ballot campaigns, the payoff is likely to be much lower. Due to the number of people these campaigns are trying to communicate with, marginal increases in effective outreach on the scale of 1% are unlikely to make a big dent in their numbers. These choices are not a failure to adopt, but a logically considered strategy not to do so.

That field teams have adopted these practices shows that the problem is not that campaigns avoid or overlook data, but that its adoption relies on being useful under constrained resources. With volunteers able to canvas, there is willing and able volunteer labor that can immediately put the data-driven targeting choices to work without much additional cost to the campaign. Moreover, as face-to-face voter contacts remain one of the most effective tactics for getting people to the voting booth, the payoff is much greater.

These findings not only call into question assumption about strategy adoption, but also about broader democratic implications of such strategies. Political communication has long been concerned about the potential for microtargeting in campaign communication—from individualized ads playing before YouTube videos, emails highlighting policies we are sure to approve of, to improved targeting of “traditional” media such as television ads. The concern that these messages can place people in a filter bubble, isolating them from important political content that would otherwise shape their views, then leads to concerns about the loss of true public deliberation and debate, increased polarization, and the development of limited versions of citizenship, wherein individuals are seen as little more than manipulable combinations of various data markers. While we should be critical of the data-driven choices and algorithms that impact our lives constantly now, we must also acknowledge that there are many areas of political messaging for which microtargeting is not of great strategic benefit, and that as a result, many of the messages we receive still follow a mass media.

Finally, just as Leticia Bode’s chapter in this volume discusses how the use of big datasets can be beneficial to qualitative researchers, so too can qualitative work about how and why big data is used give insight into the role of big data in the field of political communication. By treating big data as the object of analysis, rather than a methodological choice, this chapter reveals that campaigns’ use of data and analytics, while increasing, is not as widespread as popular press accounts may imply, and it may not be of great strategic benefit to many campaigns. Simply because it is there, and people can use it does not mean they will. Moreover, while big datasets are enticing to study because they can reveal political behaviors and trends in content on a large scale, those inquiries do not get to core questions about why campaigns create that content or citizens engage with it in a particular way. Simply because big data is there for the picking—campaigns’ Twitter feeds and Facebook posts can be scraped and catalogued—does not mean it should be the focal point of what we ought to study. Combining the use of big data with

qualitative methods investigating the actions, purposes, dispositions, and values behind such data and its strategic uses provides a holistic approach to the field of political communication as we move forward in a field that needs to keep up with rapid changes in communication technology.

## Note

- 1 Facebook had specific deals with companies to do site-wide advertising prior to this.

## References

- Baldwin-Philippi, J. (2015). *Using technology, building democracy: Digital campaigning and the construction of citizenship*. New York, NY: Oxford University Press.
- Baldwin-Philippi, J. (2016). The cult(ure) of analytics in 2014. In J. A. Hendricks & D. Schill (Eds.), *Communication and midterm elections: Media, message, and mobilization*. New York, NY: Palgrave Macmillan.
- Berelson, B., Lazarsfeld, P., & McPhee, W. (1954). *Voting: A study of opinion formation in a presidential campaign*. Chicago, IL: University of Chicago Press.
- Bimber, B. (2003). *Information and American democracy: Technology in the evolution of political power*. New York, NY: Cambridge University Press.
- Coleman, B. (2016). Let's get lost: Poetic city meets data city. In Gordon, Eric & P. Mihailidis (Eds.), *Civic media: technology, design, practice*. Cambridge, MA: MIT Press.
- Davidson, C. (2015, June). *You are not your data: What we actually do with your data*. Presented at the Personal Democracy Forum, New York, NY. Retrieved from <https://personaldemocracy.com/media/you-are-not-target-what-we-actually-do-your-data>
- Delany, C. (2011). *Online politics 101: The tools and tactics of digital political advocacy*. Colin Delany (self published).
- Delany, C. (2013). *How to use the internet to win in 2014: A comprehensive guide to online politics for campaigns & advocates*. Epolitics.com.
- Foot, K., & Schneider, S. (2006). *Web campaigning*. Cambridge, MA: MIT Press.
- Garrett, R. K., Carnahan, D., & Lynch, E. K. (2011). A turn toward avoidance? Selective exposure to online political information, 2004–2008. *Political Behavior*, 35(1), 113–134.
- Gerber, A. S., & Green, D. P. (2000). The effect of a nonpartisan get-out-the-vote drive: An experimental study of leafletting. *The Journal of Politics*, 62(3), 846–857.
- Gerber, A. S., & Green, D. P. (n.d.). Field experiments on voter mobilization: An overview of a burgeoning literature. In *Handbook of field experiments*. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab (J-PAL). Retrieved from [www.povertyactionlab.org/handbook-field-experiments](http://www.povertyactionlab.org/handbook-field-experiments)
- Green, J., & Issenberg, S. (2016, October 27). Inside the Trump bunker, with days to go. *Bloomberg BusinessWeek*. Retrieved from [www.bloomberg.com/news/articles/2016-10-27/inside-the-trump-bunker-with-12-days-to-go](http://www.bloomberg.com/news/articles/2016-10-27/inside-the-trump-bunker-with-12-days-to-go)
- Hamby, P. (2013). *Did Twitter kill the boys on the bus? Searching for a better way to cover a campaign*. Cambridge, MA: Shorenstein Center on Media, Politics, and Public Policy.
- Hersh, E. (2015). *Hacking the electorate: How campaigns perceive voters*. New York, NY: Cambridge University Press.
- Howard, P. (2006). *New media and the managed citizen*. New York, NY: Cambridge University Press.

- Issenberg, S. (2010, October 31). Nudge the vote: How behavioral science is remaking politics. *The New York Times*. New York, NY. Retrieved from [www.nytimes.com/2010/10/31/magazine/31politics-t.html](http://www.nytimes.com/2010/10/31/magazine/31politics-t.html)
- Issenberg, S. (2012). *The victory lab: The secret science of winning campaigns*. New York, NY: Crown Books.
- Judd, N. (2011, October 22). Why campaigns are happy your vote isn't as private as many think it is. *TechPresident*. Retrieved from <http://techpresident.com/news/23032/do-you-care-if-obama-knows-you-voted-what-about-if-he-told-your-friends>
- Kaye, K. (2016, June 23). Trump's first fundraising email had a 60% spam rate. Retrieved November 1, 2016, from <http://adage.com/article/campaign-trail/trump-s-fund-raising-email-a-60-spam-rate/304673/>
- Klein, D. (1989, April 16). You are where you live : Neighborhoods by the numbers: How ZIP codes can pigeonhole people like widgets in a digit factory. *Los Angeles Times*. Retrieved from [http://articles.latimes.com/1989-04-16/news/vw-2346\\_1\\_zip-code-clusters-smithsonian-magazine](http://articles.latimes.com/1989-04-16/news/vw-2346_1_zip-code-clusters-smithsonian-magazine)
- Kreiss, D. (2012). Acting in the public sphere: The 2008 Obama campaign's strategic use of new media to shape narratives of the presidential race. *Media, Movements, and Political Change: Research in Social Movements, Conflicts and Change*, 33, 195-223.
- Kreiss, D. (2014). Seizing the moment: The presidential campaigns' use of Twitter during the 2012 electoral cycle. *New Media & Society*.
- Kreiss, D. (2016). *Prototype politics: Technology-intensive campaigning and the data of democracy*. New York, NY: Oxford University Press.
- Kreiss, D., & Jasinski, C. (2016). The tech industry meets presidential politics: Explaining the Democratic Party's technological advantage in electoral campaigning, 2004-2012. *Political Communication*, 33(4), 544-562.
- Lazarsfeld, P., Berelson, B., & Gaudet, H. (1948). *The people's choice: How the voter makes up his mind*. New York, NY: Columbia University Press.
- Madrigal, A. (2013). When the nerds go marching in. *The Atlantic* (November 16).
- Malhotra, N., Michelson, M. R., Rogers, T., & Valenzuela, A. A. (2011). Text messages as mobilization tools: The conditional effect of habitual voting and election salience. *American Politics Research*, 39(4), 664-681.
- McGregor, S., Kreiss, Daniel, & Lawrence, R. (2016). Instastyle: Campaign communication in the selfie era. Presented at the American Political Science Association, Philadelphia, PA.
- Michelson, M. R., & Nickerson, D. (2011). Voter mobilization. In J. Druckman, D. P. Green, J. H. Kuklinski, & A. Lupia (Eds.), *Cambridge handbook of experimental political science* (pp. 228-242). New York, NY: Cambridge University Press.
- Nielsen, R. K. (2012). *Ground wars: Personalized communication in political campaigns*. Princeton, NJ: Princeton University Press.
- Pariser, E. (2011). *Filter bubble: How the new personalized web is changing what we read and how we think*. New York, NY: Penguin Books.
- Samuelsohn, D., & Vogel, K. P. (2016, June 28). Trump's secret data reversal. *Politico*. Retrieved from <http://politi.co/28YAt3B>
- Schwartz, R. (2009). *Cutting edge evidence-based best practices*. Presented at the Netroots Nation, Pittsburgh, PA. Retrieved from [www.youtube.com/watch?v=HE7mgWHnI0g](http://www.youtube.com/watch?v=HE7mgWHnI0g)
- Serazio, M. (2015). Qualitative political communication | Managing the digital news cycle: Power, participation, and political production strategies. *International Journal of Communication*, 9(0), 19.

- Stromer-Galley, J. (2014). *Presidential campaigning in the internet age*. New York, NY: Oxford University Press.
- Stroud, N. J. (2011). *Niche news: The politics of news choice*. New York, NY: Oxford University Press.
- Sunstein, C. (2007). *Republic 2.0*. Princeton, NJ: Princeton University Press.
- Vaccari, C. (2008). From the air to the ground: The internet in the 2004 US presidential campaign. *New Media & Society*, 10(4), 647-665.

PROOF