SIVA VAIDHYANATHAN

ANTISOCIAL MEDIA

How Facebook Disconnects Us and Undermines Democracy
What would democracy look like if Facebook’s algorithms governed the art and science of persuasion?

**PSYCHOGRAPHICS, QU’EST-CE QUE C’EST?**

A prospective villain quickly rose from obscurity, and it seemed to explain both the Brexit and the Trump upsets. On September 27, 2016, a man named Alexander Nix gave a presentation called “The Power of Big Data and Psychographics.” Nix was the chief executive officer of a market research firm called Cambridge Analytica, a part of the larger SCL Group, an American-based company owned by billionaire investor and computer scientist Robert Mercer. The board of Cambridge Analytica included one of Mercer’s friends, Steve Bannon. Bannon left the board in the summer of 2016 to take over the management of Donald Trump’s failing campaign. And Cambridge Analytica staff members were involved with the campaign to convince U.K. voters to choose to leave the European Union.

“It’s my privilege to speak to you today about the power of big data and psychographics in the electoral process,” Nix told the audience of the Concordia Summit, a European global affairs forum that encourages limited government and private-public partnerships. At the urging of Mercer, a little-known and little-liked U.S. senator from Texas, Ted Cruz, had hired Cambridge Analytica to consult on his effort to win the Republican presidential nomination.5

Cruz, Nix explained to the audience, had little chance of competing against a field of better-known opponents such as Jeb Bush, the brother of one former president and the son of another, and reality television star and failed casino owner Donald Trump. That Cruz stayed in the field of contenders for many months longer than Bush or almost all of the other challengers to Trump spoke to his use of Cambridge Analytica’s data sets and the precisely targeted advice the company gave him, Nix claimed. Cruz’s campaign had survived more than a year, until he ran out of money after losing the Indiana primary to Trump in May 2016. Ultimately Cruz won the second-largest number of committed delegates in his unsuccessful effort to wrest the nomination from Trump. After Cruz retired his campaign

Bannon convinced Mercer to support Trump, so Trump’s San Antonio, Texas–based digital team quickly made room for new partners from Cambridge Analytica.6

“Most communication companies today still segregate their audiences by demographics and geographics,” but those marks of identity only roughly predict a person’s opinions about the world, including products and politics, Nix said. “But equally important—or probably more important—are psychographics, that is, an understanding of your personality.” With the use of psychographic profiling, a marketer or campaign could precisely address an individual even if he or she stands out from the larger groups to which she belongs. So psychographic profiling uses character designs such as “openness” (how welcoming a person is to new experiences), “conscientiousness” (how much one prefers order and regularity or change and fluidity), “extraversion” (how social a person is), “agreeableness” (one’s willingness to put other people’s needs above her own), and “neuroticism” (how much a person worries). This is known in the trade as the “OCEAN” model.

Then Nix made a bold claim, one that he has not been able to support: “By having hundreds and hundreds of thousands of Americans undertake this survey (of personality traits) we were able to form a model to predict the personality of every single adult in the United States of America.”

This boast was stunning in its breadth and scope, and alarming in its potential. If this claim was true and psychographic data could be used to reliably predict the precise political inclinations of individual Americans, then a campaign could manipulate voters based on single or narrow issues, prejudices, or a lack of information. In a close election such power could sway the result if a few thousand people who would otherwise vote for a particular candidate or issue could be persuaded either to switch votes or — just as valuable to a campaign — not vote at all.

“For a primary, the Second Amendment might be a popular issue amongst the electorate,” Nix explained, referring to the provision of the U.S. Constitution that grants American broad rights to own firearms. “If you know the personality of the people you are targeting you can nuance your message to resonate more effectively with those key audience groups.”
A different collection of personality traits might demand a different sort of advertisement, Nix said. So some voters might be moved by a warm and family-oriented video that reminds a voter about the pleasures of hunting with a grandchild, for instance. Some voters need to be nudged to the left to support a particular candidate, while others might need to be nudged to the right to support the same candidate. With enough data and subtle psychographic profiles, Nix explained, a firm or campaign could develop just the right message for a particular voter or narrow set of voters.8

Again, the implications of this, if Nix’s claims are true, are substantial. They raise some serious questions. What possible platform could deliver dozens or hundreds of precisely targeted ads to individuals? Television and radio only broadcast. A campaign could use one video advertisement in New York and a different one in Texas, or one in Dallas and a different one in Houston, but getting more granular than that would be impossible. Radio allows more precision and ads are cheaper to produce, as some people congregate around certain genres of music or news. Newspapers and magazines are slow and static, offering limited ability to use narrative or hyperbole. The only platform that could deliver powerfully manipulative text or video ads so precisely to almost every potential voter is, of course, Facebook.

Just how did Cambridge Analytica gather all the personality data that filled its system? Much of it is available for sale from private data aggregators that have served marketing firms for decades. These vendors have dossiers on millions of consumers around the world, based on their purchasing records and demographic features. If Cambridge Analytica had only data from those private sources and some publicly generated data such as voter registration and voting history, Nix would be making weakly supported but bold claims about standard and well-developed techniques. His almost revolutionary boasts would not hold up to scrutiny.

The day after Trump declared victory, Nix made another bold claim. “We are thrilled that our revolutionary approach to data-driven communication has played such an integral part in President-elect Trump’s extraordinary win,” Nix wrote in a press release from Cambridge Analytica. “It demonstrates the huge impact that the right blend of cutting-edge data science, new technologies, and sophisticated communication strategies can have.” While the press release did not include the word “psychographics,” to anyone who had followed Nix’s recent speeches and his company’s embrace of psychometrics, it was easy to assume that “our revolutionary approach” meant that Cambridge Analytica had done for the winning Trump campaign what Nix had claimed it had done successfully for the losing Cruz campaign.9

Within a month, some people who followed the intersection of data and politics would make that connection explicitly, even if Nix did so only obliquely. In December 2016 an article appeared on the Swiss website Das Magazin. It generated some interest in Europe. But interest in it spiked six weeks later when the U.S.-based website Motherboard published an English-language version called “The Data That Turned the World Upside Down.”10

The article opens with an account of a young researcher named Michal Kosinski, who started his career at the University of Cambridge conducting research on psychometrics. Psychologists had forged the area of study in the 1980s and had generated the five personality traits that supplied the initials to OCEAN. Before this century a researcher could tag a person on the OCEAN matrix only after the person agreed to sit for a long questionnaire. That meant that the potential application of psychometrics was limited to those who agreed to participate in studies. From a sample of data, researchers could generate predictive models. But there was for a long time a shortage of good data to feed into and thus test and refine models. That all changed when Kosinski thought about Facebook.11

On Facebook, users were more than willing to take “personality quizzes.” These seemed harmless and fun. Tabloids had long run attractive stories about how certain preferences or behaviors “revealed your personality.” And there was a major personality testing industry, albeit specious and lacking any empirical support, devoted to personality testing to help employees perform, managers to manage, and recruiters to recruit—the Myers-Briggs system.12

Kosinski realized that he could create an application that would run within Facebook that Facebook users would gladly share. They would opt
in to letting Kosinski scrape their record of likes from Facebook as well as the answers to the test. Kosinski could then correlate the record of likes with the answers that millions of users were giving to his personality quiz. Through this method, which Facebook now forbids, Kosinski was able to generate predictive models that could indicate many aspects of identity beyond the OCEAN scale scores. Finally, a psychometric laboratory had more data than it ever imagined it would have. People volunteered these data, although it’s not at all clear that those who took the quiz fully understood the implications of allowing such deep surveillance of their social and political interactions. Still, the data came from a more “natural” setting, someone sitting at a computer or staring at a phone in an office, or on a bus, not in some contrived setting such as a university office or classroom. The model proved remarkably effective at predicting attributes. “The model correctly discriminates between homosexual and heterosexual men in 88% of cases, African Americans and Caucasian Americans in 95% of cases, and between Democrat and Republican in 85% of cases,” Kosinski and his co-authors wrote in the paper they published in 2013.13

After summarizing Kosinski’s research, the article in Das Magazin then describes an uncomfortable incident. Once of Kosinski’s colleagues at Cambridge approached Kosinski about licensing the quiz and model to SCL, which owns Cambridge Analytica. Once Kosinski realized SCL was in the business of political consulting, he refused to engage in any such collaboration or licensing agreement.14

Kosinski discovered after the surprise Brexit vote that Cambridge Analytica had boasted about using data from Facebook and the OCEAN scale to generate a model that could predict personality traits for millions of voters. The article in Dan Magazin does not state or imply that Cambridge Analytica took and used Kosinski’s quiz, model, or data. We found out in early 2018 that Cambridge Analytica relied on a different University of Cambridge researcher, Aleksandr Kogan, for that Facebook data.

After months of investigating the sources of Cambridge Analytica data, the Observer and the New York Times released reports simultaneously in March 2018 that Kogan had, in fact, given data on more than 87 million American voters to Cambridge Analytica. Reporters had found a former Cambridge Analytica data engineer who had developed misgivings about the role and purpose of the firm. “They want to fight a culture war in America,” said Christopher Wylie, the engineer-turned-whistleblower. “Cambridge Analytica was supposed to be the arsenal of weapons to fight that culture war.” Wylie revealed that Kogan had copied the user data from Facebook under the guise of performing academic research, but had sold access to the data to Cambridge Analytica. The company had built models to predict voter behavior and had convinced campaigns in the United States and around the world that the models would help target and persuade voters. When the news broke of the breadth and depth of data that Facebook had allowed out, and the fact that Facebook took no effective measures to punish companies that exploited data like that, political and commercial pressure on Facebook built to a level the company had never before experienced or expected. A movement began on Twitter urging Americans to delete their Facebook accounts. Legislators and regulators in Europe and North America launched investigations into Facebook and its data practices. And Cambridge Analytica was further exposed as a bad actor both in the methods that its leaders claimed to use and in the utter ineffectiveness of the company’s efforts on behalf of campaigns.15

**VAPORWARE EVERYWHERE**

Almost immediately after the Das Magazin article appeared in English, critics of the article and of Cambridge Analytica spoke up to undermine the association of psychometrics with electoral success. Perhaps most damning, officials who ran voter targeting and contact programs for the Cruz campaign bluntly dismissed the scheme as useless. Cambridge Analytica’s data and advice were so bad that the firm was wrong about identifying Republican voters as Cruz supporters about half the time. The Cruz campaign stopped using Cambridge Analytica after it lost the South Carolina primary to Donald Trump on February 16, 2016—three months before the Cruz campaign closed up for good. And when the Trump digital team tried to use Cambridge Analytica data, it found the older, more basic data sets offered by the Republican Party to be more reliable and useful.
Much of that standard Republican Party data derived from publicly available voter records and responses the party had gathered from voters over the previous three years.\(^\text{16}\)

As political scientist David Karpf wrote in the aftermath of the psychographic moral panic, “Targeted advertising based on psychometrics is conceptually quite simple and practically very complicated. And there is no evidence that Cambridge Analytica has solved the practical challenges of applying psychometrics to voter behavior.” In addition, the concept that Nix describes in his talk from October 2016 implies that any campaign can and would generate hundreds or thousands of tailored pieces of campaign advertising to match every combination of psychometric labels. Such a campaign would require a creative team of hundreds of writers, producers, and editors working around the clock to test various versions of an ad and quickly swapping it out for some voters and not for others. This could be done among a small number of roughly targeted set of voters—those who prioritize gun rights, others set on stopping abortion, still others who wish to stop immigration, and so on. But that targeting can and has been done for more than a decade using voter data compiled by both major parties.

“The simple explanation here is that Cambridge Analytica has been engaging in the time-honored Silicon Valley tradition of developing a minimum viable product (vaporware, essentially), marketing the hell out of it to drum up customers, and then delivering a much more mundane-but-workable product,” Karpf wrote. “The difference here is that [Cambridge Analytica’s] marketing has gotten caught up in our collective search for the secret formula that put Donald Trump in the White House.”\(^\text{17}\)

Soon even Nix and Cambridge Analytica backed down on previous claims. Many recent statements by Cambridge Analytica clearly state that the company did not engage in psychographic profiling for the Trump campaign. And for some reason the company spokespeople now insist they did nothing to help the Brexit Leave campaign. To the Trump campaign, Nix now insists, the company provided more conventional consulting advice and data analysis. There is not much evidence that Trump campaign officials appreciated even that more conventional advice from Cambridge Analytica.

This did not stop some people from continuing to draw the specter of an evil cabal of Mercer, Bannon, and Trump using our personal data and personalities to steal an Electoral College victory and vacate the will of the American voting public, which overwhelmingly supported Clinton. The New Yorker ran a story by Jane Mayer in late March 2017 recounting the story from Das Magazin. Mayer’s story was a deep and deft account of Robert Mercer’s rise as a factor in extreme right-wing politics. And Mayer mentioned Karpf’s debunking of Das Magazin. But the moment the claims of psychographic voter targeting appeared in one of the most respected magazines in the world, the issue took on new life among readers who might not follow political scientists on Twitter.\(^\text{18}\)

Even Hillary Clinton fell for the irresistible story of Cambridge Analytica practicing the dark arts of psychometrics on American voters. In May 2017 Clinton gave one of her most frank and revealing interviews after the election debacle. Speaking with technology journalists Kara Swisher and Walter Mossberg, Clinton said, “I take responsibility for every decision I made. But that is not why I lost.” Clinton told the journalists that her campaign tried to replicate the data tools developed and used so successfully by Obama in 2012. “The other side was using content that was just flat-out false, and delivering it a very personalized way, both above the radar screen and below.” Clinton explained that the Republicans had reacted to their 2012 loss by upgrading their data infrastructure so that they had drawn even with the Democrats and perhaps surpassed them. “Then you’ve got Cambridge Analytica,” Clinton said.\(^\text{19}\)

In England the Cambridge Analytica—psychometrics story also refused to die and, in fact, grew in visibility. On March 4, 2017, the Observer ran the first of what would be several stories linking psychographic voter targeting and Cambridge Analytica to the Leave campaign. The article cited a February 2016 article by Nix in the election-professional trade magazine, Campaign. “Recently, Cambridge Analytica has teamed up with Leave. EU—the UK’s largest group advocating for a British exit (or ‘Brexit’) from the European Union—to help them better understand and communicate with UK voters,” Nix wrote. “We have already helped supercharge Leave.
EU’s social media campaign by ensuring the right messages are getting to the right voters online, and the campaign’s Facebook page is growing in support to the tune of about 3,000 people per day.20

A dogged reporter and essayist, Carole Cadwalladr, has since followed up with a series of articles for the Observer that describes the growing influence of large pools of data on politics. Cadwalladr draws the same connections among Robert Mercer, Cambridge Analytica, the Leave campaign, and the Trump campaign that others have. Cadwalladr justifiably raises serious questions about the legality of potential “in-kind” contributions of consulting services by Cambridge Analytica to the Leave campaign. And she has been a fervent critic of the data industry that feeds so much personal information to those who would manipulate government for their own ends. Reporter Jamie Bartlett of the BBC has been similarly enchanted by the connections between Cambridge Analytica and the Leave and Trump campaigns. He has produced long video reports about Silicon Valley and data and how they affect our lives. In one segment Bartlett showed the empty San Antonio offices of Trump’s digital team with the side office where Cambridge Analytica staff worked. But neither Bartlett nor Cadwalladr can offer evidence that psychographic targeting per se works for campaigns. In June 2017 BBC Newsnight reporter Robert Gatehouse presented an in-depth segment on the same issues, leading off the report with spooky silent film images of hypnosis to imply that there is some sort of mind control at work in British politics. While Gatehouse’s report is ultimately fairly blunt in his dismissal of the efficacy of psychographics, it does open, close, and center the segment on the dark arts of psychographics. It’s just too good a hook for reporters to resist. All of this journalistic work has value. And because of it the UK Information Commissioner’s Office launched an investigation in 2017 into the use of private data by the Leave campaign, including its possible links to Cambridge Analytica. But by invoking psychographic profiling and manipulation as powerful and this frightening, Gatehouse, Bartlett, and Cadwalladr obscure the very real problems with the use of Big Data in politics and governance.21

THE DAMAGE

It’s not clear that Mercer, Nix, SCL, and Cambridge Analytica have succeeded in generating useful models at all, despite all the boasting Nix has done about them in an effort to seed new business. It’s also clear that by 2018 there had been no evidence that psychographics performed any better than the more standard data-intensive techniques used by the Obama and Romney campaigns of 2012 or the Clinton campaign of 2016. Leave and Trump victories were troubling and perplexing to elites and analysts. So many desperately wish to identify a magic bullet to explain complex systems generating slim differences. There were no magic bullets. There were many forces at work for both Leave and Trump that were largely unacknowledged before those votes were cast. And there was Facebook. Still, Cambridge Analytica captured all the attention, making public issues concerning Facebook, data, surveillance, and politics that had been boiling in scholarly circles for a decade.

It sounded like the stuff of spy novels. A secretive company owned by a reclusive genius billionaire taps into sensitive data gathered by a University of Cambridge researcher. The company then works to help elect an ultranationalist presidential candidate who admires Russian president Vladimir Putin. Oh, and that Cambridge researcher, Kogan, worked briefly for St. Petersburg State University. And his research was designed to develop ways to psychologically profile and manipulate voters. Before we go too deep down the rabbit hole, let’s recognize that the data Cambridge Analytica gathered to try to target more than 87 million Facebook users in the United States was not stolen from Facebook or removed after some security flaw or “data breach.” The real story is far less dramatic but much more important. It’s such an old story that the U.S. Federal Trade Commission investigated and punished Facebook back in 2011. It’s such a deep story that social media researchers have been warning about such exploitative practices since at least 2010, and many of us complained when the Obama campaign in 2012 used the same kinds of data that Cambridge Analytica coveted. Obama targeted voters and potential supporters using software that ran outside of Facebook. It was a problem then. It’s a problem now. But back in
2012, the Obama story was one of hope continued, and his campaign’s tech-savvy ways were the subject of “gee whiz” admiration. So academic critics’ concerns fell silent. Just as important, Facebook in 2012 was coming off a peak reputational moment. Facebook usage kept growing globally, as did the glowing if misleading accounts of its potential to improve the world after the 2011 revolution in Egypt. Between about 2010 and 2015, Facebook was a data-exporting machine. Facebook gave data—profiles of users who agreed to take one of those annoying quizzes that proliferated around Facebook between 2010 and 2015, but also records of those who were Facebook Friends with those users—to application developers who built cute and clever functions into Facebook. These included games like Mafia Wars, Words with Friends, or Farmville. You might have played, and thus unwittingly permitted the export of data about you and your Friends to other companies. Until 2015 it was Facebook policy and practice to let application developers tap into sensitive user data as long as users consented to let those applications use their data. Facebook users were never clearly informed that their Friends’ data might also flow out of Facebook or that subsequent parties, like Cambridge Analytica, might reasonably get hold of the data and use it however they wished.

The Federal Trade Commission saw this as a problem. In 2011 the agency released a report after an investigation revealed that Facebook had deceived its users over how personal data was being shared and used. Among other violations of user trust, the commission found that Facebook had promised users that third-party apps like Farmville would have access only to the information that they needed to operate. In fact, the apps could access nearly all of users’ personal data—data the apps didn’t need. While Facebook had long told users they could restrict sharing of data to limited audiences like “Friends Only,” selecting “Friends Only” did not limit third-party applications from vacuuming up records of interactions with Friends. The conclusions were damning. They should have alarmed Americans—and Congress—that this once huggable company had lied to them and exploited them. Through a consent decree with the commission, Facebook was barred from making misrepresentations about the privacy or security of consumers’ personal information. It was required to obtain consumers’ affirmative express consent before over-riding privacy preferences. And Facebook was required to prevent anyone from accessing a user’s material more than thirty days after the user has deleted his or her account. Most important, Facebook had to proactively police its application partners and its own products to put user privacy first. The consent decree put the burden on Facebook to police third parties like Kogan, the Obama campaign, and the makers of Farmville. Facebook was responsible for making sure fourth parties, like Cambridge Analytica, did not get and use people’s information. We now know how well Facebook lived up to that responsibility. Facebook shut down this “Friends” data-sharing practice in 2015, long after it got in trouble for misleading users but before the 2016 election got into high gear. Not coincidentally, Facebook began embedding consultants inside major campaigns around the world.

For 2016 Facebook would do the voter targeting itself. Facebook is the hot new political consultant because it controls all the valuable data about voter preferences and behavior. No one needs Cambridge Analytica or the Obama 2012 app if Facebook will do all the targeting work and do it better. This is the main reason we should stay steady at the rim of the Cambridge Analytica rabbit hole. Cambridge Analytica sells snake oil. No campaign has embraced it as effective. Cambridge Analytica CEO Alexander Nix even admitted that the Trump campaign did not deploy psychometric profiling. Why would it? It had Facebook to do the dirty work for it. Cambridge Analytica tries to come off as a band of data wizards. But they are simple street magicians, hoping to fool another mark and cash another check.22

We should be wary of the practice of data-driven voter targeting in general—whether done for the Trumps of the world or for the Obamas of the world. The industry devoted to rich data targeting and voter manipulation is far bigger than SCL and Cambridge Analytica. It’s growing on every continent. And it’s undermining democracy everywhere. Facebook is doing the data analysis internally. Facebook is working directly with campaigns—many of which support authoritarian and nationalist candidates. You don’t need Cambridge Analytica if you have Facebook. The impact of Facebook on democracy is corrosive.23
Chapter 6: The Political Machine


4. "2016 Election Results," "Presidential Election Results."


Chapter 7: The Disinformation Machine


