Everything you need to know about the Cambridge Analytica-Facebook debacle

PHILIP BUMP MARCH 19, 2018

Late on Friday, Facebook made an unexpected announcement: The data firm Cambridge Analytica, hyped as integral to President Trump's election, was suspended from the social network for using data collected improperly from Facebook users.

It is a complicated issue that many people might have missed, given the timing of the announcement. With that in mind, here is an overview of the groups involved, what happened — and what it means.

1. What is Cambridge Analytica?

Cambridge Analytica is a data firm that promises its customers insights into consumer or voter behavior.

On the commercial side, that means tools like "audience segmentation" — breaking out advertising audiences into smaller groups — and then targeting advertisements to those groups on "multiple platforms."

On the political side, it is much the same thing, with one tweak. While advertisers generally target consumers as groups, political campaigns need to target specific *people* — registered voters receptive to a potential message.

"Combining the precision of data analytics with the insights of behavioral psychology and the best of individually addressable advertising technology," the company's website pledges, "you can run a truly end-to-end campaign." And that is why Cambridge Analytica was created.

Robert Mercer is a prominent conservative donor whose public profile rose sharply over the past few years. He and his daughter Rebekah invested millions in efforts to reshape conservative politics, funding Citizens United, the anti-mainstream-media Media Research Center and Breitbart News.

In 2013, Robert Mercer partnered with a British firm called SCL Group and its elections director Alexander Nix to test SCL's methodology in Virginia's governor's race, as the New York Times reported. Their candidate, Republican Ken Cuccinelli, lost. But the Mercers moved forward with a political data strategy anyway, partnering with Nix to create

Cambridge Analytica, which would use SCL's data and methodology for political work.

2. What prompted the Facebook suspension?

psychology" line.

There are lots of data companies that can tell you who's registered to vote, and there are lots of companies which compile consumer data on those same voters. This, in fact, was an instrumental part of Facebook's sales pitch to political campaigns (back before it quietly buried that pitch in the wake of questions about Russian interference in the 2016 election). After the 2014 election, we wrote about how Facebook offered campaigns a place to overlap their voter data (who's registered and basic demographic information) with Facebook's vast array of data on its users' behavior. While most firms that collect data on consumer behavior do so by tracking the bread crumbs we leave around our consumer culture — grocery store rewards cards, magazine subscriptions, etc. — Facebook has the advantage that so many Americans *tell* the company precisely what they like, by quite literally clicking the "like" button.

Facebook's database of personal information may be the largest in the world, given that nearly a third of the globe has an account with the company. If you are a company looking to provide data services, you would justifiably by jealous of the information Facebook possesses. So Facebook (recognizing an opportunity when it sees it) provides a way for software developers to build on top of their platform, allowing other companies to use their data under certain conditions. It used to be fairly trivial, in fact, for developers to build an application that would then pull a great deal of information from the site, including information about your friends' activity. In May 2014, the site announced it was tightening that access, beginning the following year.

That change came slightly too late.

To apply its "insights of behavioral psychology" to national politics, as the Mercers intended, the SCL/Cambridge team needed a lot of information about a lot of Americans. According to the Times's report, a Cambridge employee named Christopher Wylie encountered a researcher at Cambridge University named Aleksandr Kogan. Kogan built an application that leveraged Facebook's tools to pull information from the site and then pitched its use using Amazon's Mechanical Turk, a tool that allows developers to hire humans (sometimes then referred to as "turkers") to do simple tasks for small fees.

The Intercept reported on how it worked last year.

> "The task posted by 'Global Science Research' appeared ordinary, at least on the surface.

> The company offered turkers $1 or $2 to complete an online survey. But there were a couple of additional requirements as well. First, Global Science Research was only interested in American turkers. Second, the turkers had to download a Facebook app before they could collect payment. Global Science Research said the app would 'download some information about you and your network … basic demographics and likes of categories, places, famous people, etc. from you and your friends.'"

Global Science Research was Kogan. Using this method, he gathered information on tens of millions of Americans. (The Times says more than 50 million; other outlets say 30 million.) That information was then used to build out SCL/Cambridge Analytica's profiles.

In building his Facebook application, Kogan had pledged that his data collection was only for research purposes and that it would remain anonymized — not able to be linked to specific people. When the Guardian reported in late 2015 on the link between Kogan and Cambridge, it prompted Facebook to promise to investigate the situation. (The Guardian's story was pegged to Sen. Ted Cruz's (R-Tex.) presidential campaign using Cambridge Analytica for its voter contact efforts. Cruz was strongly supported by the Mercers, who also created well-funded outside groups to promote his candidacy.)

In its statement on Friday announcing the suspensions, Facebook carefully put the blame on Kogan misusing its tools and explained it had demanded in 2015 that Kogan, SCL and Cambridge delete its Facebook data. The suspension was prompted by learning last week — apparently after being contacted by the Times — that Cambridge was still in possession of some of the Facebook data. (The company denies that.)

3. What does Cambridge Analytica's data actually look like?

It is not clear, but we do have one hint.

A professor at New York's New School named David Carroll was studying ad targeting when he realized Cambridge's link with SCL meant the company might be subject to Britain's broader data-access laws, allowing him to potentially see what data the company had collected on him. In March 2017, he got a response.

It is not clear how Cambridge/SCL developed that profile of Carroll (which he described on Twitter as "accurate"), but one can see how Facebook data might help inform those categories.

One can also see how, once the profile was developed, the Facebook data underlying it would become unnecessary. It is as though you snuck a peek at the secret recipe for Kentucky Fried Chicken and then developed your own recipe based on it. You may not be in possession of the recipe, but that is sort of beside the point.

4. Where does the Trump campaign fit into this?

Trump's digital team was run by Brad Parscale, who last month was named campaign manager for Trump's 2020 effort. Trump's general election campaign was slow to get geared up after the primary, and, by mid-2016, there was a debate over how to invest in digital marketing. Bolstered by Parscale's advocacy (and Jared Kushner's championing) the campaign hired Cambridge Analytica, over then-campaign chairman Paul Manafort's apparent objections. The decision may have been made easier, too, by Cambridge/SCL's role in the successful Brexit campaign in Britain the same month.

As noted above, the Mercers had been hoping Cruz would be the Republican nominee. Once Trump won the Republican nomination, though, they shifted their focus. (The extent to which the hiring of Cambridge Analytica greased that transition is not clear.) They were reportedly instrumental in the August 2016 overhaul of Trump's campaign, recommending the hiring of both Stephen K. Bannon (from Breitbart) and Kellyanne Conway, who had been working for one of their pro-Cruz PACs.

Over the last few months of the campaign, Parscale's team invested heavily in Facebook advertising, even hosting a Facebook employee at their Texas war room who helped guide their work. The advertising the campaign deployed was informed by Cambridge Analytica's data.

Bloomberg reported on the data team shortly before the election and how Parscale managed the competing data from Cambridge and the Republican Party.

"Parscale was building his own list of Trump supporters, beyond the RNC's reach," Bloomberg's Joshua Green and Sasha Issenberg wrote. "Cambridge Analytica's statistical models isolated likely supporters whom Parscale bombarded with ads on Facebook, while the campaign bought up email lists from the likes of Gingrich and Tea Party groups to prospect for others."

One footnote: Campaign adviser Michael Flynn also contracted with SCL shortly before the end of the campaign, though he apparently never did any work for the company.

5. Does this mean that Trump won the election unfairly?

Well, this is a broader question: Does Cambridge actually help win elections? Or, put another way: How much of Cambridge's rhetoric about psychographics is just hype?

6. Fine. Where *has* Cambridge Analytica won elections?

In most cases, it's very hard to identify one particular factor that made the difference in a

political campaign. Despite the ubiquity of politicking, campaigns don't happen that often and, when they do, there are thousands of factors that make each contest unique. So analyzing the effects of campaign tactics means perusing a small sample in which we're asked to compare apples to oranges to grapes to dogs to stars to love to six.

This is hugely advantageous for political consulting firms because it's often hard to check their claims about how effective they are. Politicians are deeply superstitious and seize on their own and others' past successes to guide their decisions moving forward. What's more, the field of data-driven political persuasion is fairly new, meaning that a company which can claim success in a realm that many career politicians don't really understand has a huge marketing advantage. Say that you have cracked the code to targeting voters with specific messages, and a lot of campaigns will write you checks.

Cambridge Analytica hasn't been around that long, but they've been involved in several successful campaigns. There was Sen. Thom Tillis's (R-N.C.) Senate campaign in 2014, which he won by 1.5 points. There was the "Leave" campaign in the United Kingdom in 2016 which won by 3.8 points. And there was Trump, who lost the popular vote by 2.1 points but who won the electoral college.

There were also losing campaigns. Before Trump, the highest-profile effort Cambridge undertook was Cruz's — and he lost. Sure, he ended up in second place in the delegate count despite being fairly unpopular the year before, but his strategy was like Trump's: leverage a core base of support to ride out a crowded field of candidates.

In June 2016, Politico reported that Cruz's team "was disappointed in Cambridge Analytica's services and stopped using them before the Nevada GOP caucuses in late February, according to a former staffer for the Texas Republican."

So it's hard to say in the abstract the effect that Cambridge might have had in Trump's race — and that it's harder still to say what role the laundered Facebook data played.

Two days before the election, Cambridge's Nix said in an interview that his firm actually *wasn't* able to leverage its psychographics on Trump's behalf.

Here's Nix, speaking to TechCrunch:

> "We just didn't have the time to rollout that survey. I mean, Christ, we had to build all the IT, all the infrastructure. There was nothing. There was 30 people on his campaign. Thirty. Even Walker it had 160 (it's probably why he went bust). And he was the first to crash out. So as I've said to other of your [journalist] colleagues, clearly there's psychographic data that's baked-in to legacy models that we built before, because we're not reinventing the wheel. [We've been] using models that are based on models, that are based on models,

> and we've been building these models for nearly four years. And all of those models had psychographics in them. But did we go out and rollout a long form quantitative psychographics survey specifically for Trump supporters? No. We just didn't have time. We just couldn't do that."

An important asterisk: Two days before the election, Nix (and nearly everyone else in America) likely thought Trump was going to lose. A good way for a political consulting company to cover its back in the event of a loss is to say that it didn't have the time to deploy its core value proposition.

7. Is special counsel Robert Mueller tracking this whole thing?

Apparently.

Given that the Trump campaign and Cambridge invested so much in targeting people online and given that we know that Russian actors tried to leverage Facebook ads and social media to influence voters, there's a natural question as to whether those two efforts had any coordination.

In July, McClatchy reported that Mueller's team was looking specifically at that.

"Congressional and Justice Department investigators are focusing on whether Trump's campaign pointed Russian cyber operatives to certain voting jurisdictions in key states," Peter Stone and Greg Gordon wrote. They quoted a former Pentagon staffer named Mike Carpenter. "There appears to have been significant cooperation between Russia's online propaganda machine and individuals in the United States who were knowledgeable about where to target the disinformation," Carpenter said.

8. So are there links to Russia?

Well, it depends on what you mean by "links." We're in this weird moment where any even tangential link to Russia or a Russian person is heralded as a sign of questionable collusion.

So here's what we know.

The Times reports that SCL Group had spoken with the Russian oil giant Lukoil in 2014 and 2015, and that the company "was interested in how data was used to target American voters, according to two former company insiders who said there were at least three meetings with Lukoil executives in London and Turkey." (In an interview with the "Today" show on Monday, Wylie reiterated this claim.)

The paper also notes that Cambridge included questions about Russian president Vladimir

Putin in 2014 focus groups, though we'll note that this was also the time period in which Russia's seizing of Crimea became central to American foreign policy conversations.

Late last year, it was reported by the Daily Beast that Nix had contacted WikiLeaks's Julian Assange before the election offering to host emails stolen from Hillary Clinton's campaign chairman to create a searchable database. Assange declined the offer. Those emails are believed to have been stolen by Russian hackers linked to the country's intelligence agencies.

One other link is worth mentioning. Kogan, the Cambridge researcher who developed the tool that led to the Facebook suspension, had reportedly also received a grant from the Russian government to research social media.

"Nothing I did on the Russian project was at all related to Cambridge Analytica in any way," Kogan told the Guardian.