CARNEGIE COUNCIL

#Republic: Divided Democracy in the Age of Social Media

May 4, 2017

Detail from book cover

Cass R. Sunstein

Joanne J. Myers

JOANNE MYERS: Good morning. I'm Joanne Myers, director of Public Affairs Programs, and on behalf of the Carnegie Council I want to thank you all for beginning your day with us.

Our speaker, Cass Sunstein, is one of our country's most influential legal scholars. Each time Professor Sunstein has spoken here he has encouraged us to think about the importance of dissent and participation for sustaining a healthy and robust democracy. I am confident that this morning he will inspire us to do so once again.

Shortly he will be discussing his most recent book, entitled #Republic: Divided Democracy in the Age of Social Media. In it he writes about the critical relationship between democracy and the dangers that the Internet poses for our politics. For further discussions on this topic you can visit our website at www.carnegiecouncil.org and type in "Cass Sunstein" and/or "democracy."

It has been said that "democracy is not a spectator sport." These words imply, and even insist, that action is required for democracy to flourish. Yet, in this digital age of social media there is a tendency toward complacency as it has become far too easy to search only for the facts that affirm one's point of view. The problem with customizing our informational environment is that it makes it less likely that we will serendipitously encounter perspectives different from our own and those which could ultimately change our mind. Why does this matter? Because, as Professor Sunstein states, "the lifeblood of democracy is to entertain a variety of viewpoints."

In #Republic Professor Sunstein shares his concerns that data algorithms are often being used to politically manipulate people in unprecedented ways, needs, or beliefs. At the same time there is a danger that communities of like-minded individuals will become more isolated and their views more extreme.

The big question is how to address this modern-day challenge. Are there ways to use social media, technical or otherwise, that will expose us to diverse views and ensure that deliberative democracy endures? For insight on this issue, please join me in welcoming our speaker, one of the most steadfast defenders of democracy, our guest today, Cass Sunstein.

Thank you for coming.

CASS SUNSTEIN: Thank you so much. Thank you all for coming so early in the morning.

I think everything I have to say is captured in two tales and three social science experiments. So you're going to hear five things. There are two rivulets, which are the tales that led to this project.

I was blessed—or cursed—to live in the Waldorf Towers for most of the last few years because my wife Samantha Power was U.S. ambassador to the United Nations. For a country boy from Woburn, Massachusetts—a suburban boy—it was pretty amazing to live in the Waldorf Towers. But more amazing was to be greeted every morning and every afternoon with an extremely cheerful, "Good morning, Mr. Power; good evening, Mr. Power; how was your day, Mr. Power?"

As the "Mr. Powers" mounted, there was one person, one interlocutor, with whom I actually became friends. He was one of the concierges there and became a person with whom I talked about things other than how the weather was. I thought I should actually tell him my name.

Several months in, I said, "You know, my name, actually it's Cass, Cass Sunstein. You can call me Cass or Mr. Sunstein, whatever you like, but my name is Cass Sunstein." And he looked at me and he said, "That's amazing. That's incredible. You look exactly like Mr. Power." [Laughter]

On reflection, what was noteworthy about that was that he was updating based on the prior convictions he had in a way that social scientists have studied. He was engaged in what's called Bayesian updating. Given his prior conviction that Samantha Power's husband was a little over six feet tall, just starting to lose his hair, and looked like this, it was more likely that there were twins walking around his building than that Ambassador Power's husband had a name other than Mr. Power. That exercise in information updating based on prior conviction actually tells us a lot, I think, about current debates, whether the issue involves immigration or the Affordable Care Act or terrorism.

The second tale is serendipitous given the day. It's May 4. I hope every one of you knows it is International Star Wars Day, and the predecessor to this book—and if you'll forgive me, we're going to get into Star Wars arcane matters, stay

with me, it won't last long—there was a debate between two of America's great screenwriters before the third movie, Return of the Jedi. If you don't know anything about Star Wars, that's fine. The debate is what matters, not the content of the Star Wars narrative.

Lawrence Kasdan, who is a sensational, and I think in many ways deep, screenwriter, said to George Lucas, "Luke Skywalker in the third movie has to die."

George Lucas, namesake not coincidentally of Luke, said "Luke is not going to die."

Then Kasdan said, "Okay, Princess Leia has to die."

Then George Lucas said, "You know what? Princess Leia is not going to die."

Then Kasdan said, "Okay, Yoda has to die."

Lucas says, "Kasdan, he's not going to die. I mean, maybe in a way, but he'll come back to life." He said, "It's not nice going around killing people."

At that point Kasdan got deep in terms of his own convictions about the nature of art and Kasdan said, "George, art has more meaning if someone you love is lost along the way. The work binds itself to the audience if that happens. Someone really has to die for this work to have the kind of size that you want from it."

Lucas responded very quickly, "I don't like that and I don't believe that."

Now those, in my view, are precious words. They're not like the words from the concierge at the Waldorf. They're not about updating based on your prior conviction. They're about motivated reasoning.

Notice the sequence in Lucas' words: "I don't like that and I don't believe that." Dislike precedes disbelief. That tells us a fair bit, I think, about information processing in the modern era—and also in the not-so-modern era, but most notably in the modern era—where not liking and not believing can be instantaneous on your Facebook page or Twitter feed.

Those are the two tales. One is about updating based on prior conviction; the other is about motivated reasoning.

Now I'm going to tell you about three experiments. The first is something I was engaged in a number of years ago in Colorado, and it involved two places, Colorado Springs and Boulder. Think of Boulder as like San Francisco or Cambridge or many places in Manhattan, which is to say left of center, and think of Colorado Springs as "Trump country," it's conservative.

What we did was to get groups together, five or six people, multiple groups, and ask them for their private views about climate change, same-sex unions, and affirmative action, to deliberate with people in their same town about each of those views, to reach some kind of verdict, and then to state their views anonymously and privately. Really simple experiment.

My own interest was in one aspect of the experiment only: What would happen to the private anonymous views of the people in Boulder as a result of brief conversations with other people in Boulder about the three issues; and, similarly, what would happen to the private anonymous views of the people in Colorado Springs on the three issues? As a little experiment, which we won't conduct formally, maybe you can imagine to yourself what sort of shift you would predict would happen in Boulder and what sort of shift you would expect, if any, to happen in Colorado Springs.

What happened for all of our groups was exceedingly simple, and it was the same thing. In Boulder, people got on

all three issues more unified, more confident, and more extreme. In Colorado Springs guess what happened? They got more unified, more confident, and more extreme. In their private anonymous views before they talked with like-minded others the Boulder people were probably a little bit to my left and the Colorado Springs people a little bit to my right. After our experiment they were occupying different political universes.

The old Colorado experiment is basically a fruit fly experiment version of what we're observing on social media and with web choices every minute of every day where people are, by algorithm or by their own decisions, being sorted into like-minded groups.

A question is: Why did this happen in our Colorado experiment? I've actually seen the tapes and, both, theory about the phenomenon, which is called "group polarization," and just viewing the tapes shows you the three explanations.

The first, and the most straightforward, is that if you have a group of people who think that climate change is a serious problem in Boulder, they are going to give information that supports that judgment and the amount of information that goes the other way will be pretty small, and if everyone is listening to each other after they've conversed, they will be more unified, more confident, and more extreme. It is just a statistical fact that if you have a room of people of whom the majority thinks that the immigration problem needs to be solved in a particular way, the arguments that support that judgment will dominant in numerosity the number of arguments the other way and, if people are listening, that is what is going to end up prevailing. You could see that in real time.

The second explanation is slightly more subtle, and I think the technical social science term is it's "cooler." Here's how it worked. Most people on a political issue that has a degree of complexity are humble and kind of tentative. So if you ask random people in Boulder or Colorado Springs, "Should the United States sign an international agreement on climate change?" they might have a tendency, but it is a little soft. That suggests that on a scale, let's say, of 0-to-8, where zero means "No, oh my god" and 8 means "definitely," if they're negative they'll tend to be around 3. But if someone in the room corroborates their initial conviction, then they will become clearer in it, and as their clarity increases their extremism increases as well.

In real time you can see the people in Colorado Springs on climate initially saying, "Well, maybe we shouldn't sign that treaty." Their friends or their acquaintances in the room say, "Maybe we shouldn't sign that treaty, I agree." Then they say, "Crazy to sign that treaty." That's the dynamic that the data I just described observed.

The third phenomenon, I think, helps account actually for the Arab Spring, the fall of communism, and the fall of apartheid. It's really simple. It's that if you find yourself in a group of people who are relevantly like you and they tend to think, let's say, affirmative action is a bad idea, even if you think maybe it's a good idea, to buck the trend takes a degree of courage. If the group is kind of unified in its skepticism about affirmative action, your own concern for your reputation will produce a degree of self-silencing.

That will lead people in the group verdict to end up subscribing to a position that is different from the position with which they started. Then, if they're asked anonymously what do they think, it's a little embarrassing to say, "Actually I don't think the thing I just said publicly," which means that the reputational pressure leads to a public conviction, which leads to a distortion of the anonymous statement. That's experiment 1.

Experiment 2 is, I think, more startling. It has now been corroborated with reams of data. There is nothing artificial about it. This doesn't involve lawyers and psychologists and economists engaging with people. It involves a gold mine.

The gold mine is that in the United States we have had for a long time now courts of appeals that are three-judge panels, and they consist either of three Democratic appointees (DDD), three Republican appointees (RRR), two Democratic appointees and one Republican appointee (DDR), or two Republican appointees and one Democratic

appointee (RRD). That's it. Because the draw is random and because the number of votes is in the tens of thousands, you have a lot of statistical power and you can find out what is the likelihood in an ideologically controversial case of sex discrimination, sexual orientation discrimination, the environment, rights of labor—whether the political party of the appointing president is a good predictor of how a judge is going to vote. No big news here. It's a pretty good predictor of how a judge is going to vote.

But—and this is a way into the finding I'm going to highlight here—as good, and often a better, predictor of how Judge Smith is going to vote than whether Judge Smith was appointed by Clinton or Bush, if you want to know how Judge Smith is going to vote, don't ask whether Judge Smith was appointed by Clinton or Bush. Do ask, "Who appointed the two other judges on the panel with whom Judge Smith is sitting?" Judge Smith's vote is better predicted by the political affiliation of the president who appointed the two panel colleagues than the political affiliation of the president who appointed Judge Smith.

To make this more vivid, I'll give you some concrete numbers. The distance between the liberal voting patterns of Republican and Democratic appointees across a very large data set is about 13 percent—liberal votes 38 percent from Republican appointees, 51 percent by Democratic appointees. You don't take that to the bank, but that's basically what it looks like.

If our Republican appointee is sitting with two Republican appointees and a Democrat is sitting with two Democratic appointees, that 13-point disparity often grows to a 30-to-40 percent disparity; which is to suggest on an RRR panel, if the Environmental Protection Agency is being challenged by coal companies, good luck Environment Protection Agency. In a sex discrimination case, if General Motors is being sued by a woman before a DDD panel, that woman is highly likely to win.

From the standpoint of the rule of law, this is very disturbing, yes, that the random has such power. But for present purposes I just described the Colorado experiment in a most unlikely place, the federal judiciary. DDD panels are Boulder and RRR panels are Colorado Springs. Where an issue isn't political, it involves law, which is supposed to turn on something other than political predilection, it turns out that the echo chamber effect is brutally effective in determining outcomes.

What makes this startling—and that's the word I used—is something I still have a very hard time wrapping my mind around. Imagine it's a case involving the rights of a labor union and it's a DDR panel, two Obamas and a Bush. The Ds have the votes; you only need two, you don't need that R—"Go dissent, R." If it's an Environmental Protectional Agency rule and it's an RRD panel, the Rs can rule against the EPA; they have no need for anything from the D. And yet, on an RRR panel the Republicans are far more conservative than on an RRD and on a DDD panel the Democrats are far more liberal than a DDR. That is showing the power of the echo chamber.

I only have one more experiment for you and then I'm going to talk a little bit about Facebook and Twitter. This is, I confess, my favorite of the three. It's the most recent and the data is still being collected, but I have the headlines for you here. To get at this we're going to have to do a little experiment that won't seem to be related to politics and law, but it is the actual, speaking autobiographically, inroad into the work I'm going to describe.

Can you—you don't have to write it down or tell anyone—rank yourself in terms of how good-looking you think you are on a scale of 0-to-10?

Now I have some news for you. Remember Brad Pitt and Angelina Jolie when they were at their very best-looking before they got divorced? You look like them; that's a 10. Now what do you think? Imagine I said that credibly.

It turns out that if people rank themselves as a 6 on the 0-to-10 scale and then are told something like what I just told you credibly, they will incorporate the good news. They'll go up to 8 or 9.

Now let's do it the other way. Imagine that I told you Danny DeVito and Rhea Perlman—you know who I'm talking about?—that's what you look like. I don't mean to insult those people or you. You're kind of a 2. Now what do you do? How do you change your number? If you're like most people, you will not change your number much or at all.

With respect to looks—and this has been tested robustly—people will update with good news and they will ignore bad news. Good news is credible, bad news is "what do they know?" That's about looks.

The same phenomenon—it's called the Good News-Bad News Effect—has been observed for the likelihood that people have susceptibility to diabetes, that they are going to be trapped in an elevator, that they are going to have a mouse in their house, that they are going to split from their partner, etc. The human mind is more willing to incorporate and update with good news than bad news, showing that Lucas' "I don't like that and I don't believe that" is hardwired into the human species. By "hardwired" I mean because neuroscientists have shown there is actually an identifiable part of the brain that blocks incorporation of bad news, such that the phenomenon I'm describing dissipates if you zap that part of the brain; it goes away entirely. This work done by economists and neuroscientists has intrigued me for the last four or five years, and I thought, There have to be political implications.

Here's what we've got. We divided America—I worked with neuroscientist Tali Sharot, who is, more than anyone, responsible for understanding the Good News-Bad News Effect—into three tertiles based on data in terms of climate change. Creatively, we called them "strong, moderate, and weak believers." We identified this by asking them, "Are you an environmentalist? Do you think the United States was right to sign the Paris accord? Do you think that man-made climate change is occurring?" on a scale of 0-to-7. The people at the top of the scale were strong believers, the bottom weak believers, and the middle moderate believers.

Notice, if you would, that the weak believers in climate change are not skeptics; they are just in the bottom tertile of America. They don't think that this isn't happening. The study I've just described no one would publish.

The next part I'm going to describe no one would publish either, because it's too simple, which is: Ask the people in the three tertiles: "How warm do you think the planet is going to get by 2100 if we don't do anything?" The top tertile said 6.3 °F warmer, the bottom 3.6 °F warmer, and the middle 5.9 °F warmer. No one would publish that either, because it's too straightforward. But it's pleasing that the strong believers, as measured by our index, think it is going to get hotter than the moderates, who think it is going to get hotter than the weak believers.

Now I'm going to tell you the part of the study that is the money part. We divided half of our sample into a group that got news. The news is: "Hey, scientists think that the situation is a lot better than they had thought. The warming problem is much less severe. They expect warming by 1-to-5 °F only." The other half got bad news—Danny DeVito/Rhea Perlman as opposed to Brad Pitt/Angelina Jolie—"It's going to get really hot." The other half got "Scientists think it's going to be 7-to-11 °F, much worse than had long been projected."

What we wanted to know was how would our three tertiles react to the good news and to the bad news. With respect to the bottom tertile, the 3.6, the finding is very straightforward: They were much moved by the good news; they fell from 3.6 on average to 2.6 on average, which is a whopping change because the projection wasn't zero, it was 1-to-5. Good news they updated significantly.

The bottom tertile getting the bad news, guess how much they updated? Zero. This isn't climate skeptics, this is Americans who think it is going to get hotter. But the 7-11 °F had no impact on their judgments.

In terms of the political dynamics of information updating, what I've just described to you strikes me as instructive. In terms of social science, it's basically replicating the Good News-Bad News Effect, which we've known about mouse in the house or trapped in the elevator or how good-looking are you. The bottom tertile finds that bad news is uninstructive and good news is "Oh my god."

The top tertile, what do you think would happen with them? If they are like everybody else in all the other experiments, they would show the same pattern. But they didn't. The top tertile of Americans by our measure were highly responsive to the bad news. Hearing the 7-to-11 °F estimate, their own judgment leaped from 6.3 to 8.3. That's a huge update. The good news they weren't utterly indifferent to, but it had a much more modest effect; their estimate fell by less than a degree.

Now, what you're hearing about is asymmetrical updating as it is described, but of two radically different kinds, where the top tertile asymmetrically updates in the sense that it is responsive to the bad news and the bottom tertile asymmetrically updates in the sense that it is more responsive to the good news. That is, isn't it, a kind of ridiculously crude depiction of what America is getting with respect to climate change, the impact of the Affordable Care Act, how President Trump is doing, whether [Hillary Clinton](#) is a criminal, whether immigration should be fixed one way or another? They are getting exactly the same version of the good news-bad news and asymmetrical updating should produce the kinds of polarization that we produced with our tiny little experiment.

After our little experiment, the climate change top tertile was at 8.3 and the bottom tertile was at 2.6, even though before the experiment the disparity was much smaller, 3.6 versus 6.3. The social media process, it's being iterated, it's happening over and over again. We just did a one-shot.

Why did our climate change results come out as they did? I confess I don't know, but I have two explanations, one of which is right. We're trying to figure out which one.

The simplest explanation, I think, is it's motivated reasoning. The bottom tertile is pleased to hear the Earth is not going to get so warm. That's good news for the planet and it's good news for their own non-terror about climate change. They find it affirming and they find it pleasing.

The bad news they think is like being told "you look terrible." It's unpleasant because it suggests the planet is going to burn up and it's unpleasant because it says you're fundamentally wrong in something on which you have a conviction.

The harder people to explain are the top tertile. Why are they more willing to believe the bad news than the good news? On this motivated reasoning account—I find this intriguing, I confess—they are so deeply invested in their conviction that the planet is going to get really hot that they would rather believe the planet is going to get really, really, really hot than that they're wrong.

That is humanly recognizable, isn't it, where, whether you're on the right or left, to find that the thing that worries you actually is a non-problem is not pleasing, even though it's actually very good news for America and the world? That's the George Lucas "I don't like that and I don't believe that" account.

The other one is the Waldorf, "That's amazing, that's incredible." The other account would be just, given people's prior convictions, where the strong believers think, Look, it's going to get really hot, when they see prominent scientists think otherwise, they think to themselves, Who paid for that. Exxon? It has no credibility. Whereas when they get the news that it's going to get really hot, they say, "Oh, it's even worse than I thought."

It's a little like if you read something saying "dropped objects don't fall," that's not going to move you to disbelieve in gravity. Whereas the weak believers, if they hear "prominent scientists conclude it's going to be 7-to-11," they think, That's completely ridiculous, that's environmental nuts, that's not real science, that's junk, and they dismiss it like my Waldorf guy did, and that's not irrational, given their prior conviction. If they hear actually "It's not going to get so hot," they think, Oh, that's real science. It's better than I thought.

Now let's talk about Facebook by way of conclusion. What I'm going to do is juxtapose Facebook with [Jane Jacobs](#).

Cass R. Sunstein: Facebook's Misguided "Architecture of Control"

▶

Facebook had a post in 2016—pretty recent—in which it described its News Feed and the core values that animate it. What it said is: "Imagine that you are given thousands of news stories and that you're in a position to choose 10 or a few that interest you. Shouldn't you be able to? You should get something that is unique, subjective, and personal, that's yours. That's the value of News Feed, and that's what we're trying to achieve."

Facebook is a company for which I have a lot of admiration, both because it's providing something great and because the people there are completely amazing. But the lack of self-consciousness about the celebration of the architecture of control is what I want to put a spotlight on.

What Facebook is doing in 2016 in celebrating the unique and personal and subjective is basically having kind of a bunch of balloons descend about the creation of what in the 1990s technology specialist Nicholas Negroponte from Massachusetts Institute of Technology called "the Daily Me," urging that technology would ultimately enable us to create. The Daily Me doesn't mean the main newspaper, it means the daily you, me, completely personalized—yours. And Facebook is saying that is what News Feed does.

I think the technical response to Facebook's statement of its core values is WTF—if you don't know what that means, good, let's try OMG, which is the G-rated version (Oh my god!)—in the sense that that is a conception of what News Feed should look like, that takes a stand in favor of an architecture of control which couldn't be more opposed to—this is the year of Alexander Hamilton—Hamilton's conception of self-government, which celebrated heterogeneity and diversity and the forms of jarring—that's his word, "jarring," in a good way, jarring in a productive way—that come from encounters with diverse others.

Jane Jacobs had a different vision in her prose poem The Death and Life of Great American Cities. Jacobs was writing about the architecture that New York or Berlin or Paris or San Francisco have. What she urged was that people can be acquaintances with, on familiar terms with, others who have life stories or current experiences or tragedies or joys that are completely different from their own. What Jacobs celebrated was the fact that serendipitously you might turn a corner in New York and see someone who has a job or a project or a store or a product to sell or a complaint to make, whom you never would have chosen to see, but that person might change your day, conceivably your month, and in some cases your life.

If you think to yourself, I'm going to take a little gamble here, about how you ended up where you are in life, whether it involves job or focus or friends or partners, the extent to which serendipitous encounters produce that thing tells a tale of life's narratives, and democratic self-government too, that Facebook's statement of core values, I think, can't capture. That's Jane Jacobs, our hero.

The hero of these remarks is both Jacobs and [John Stuart Mill,](#) who [wrote](#) a long time ago: "It is hardly possible to overrate the value, in the present low state of human improvement of placing human beings in contact with persons dissimilar to themselves, and with modes of thought and action unlike those with which they are familiar. Such communication has always been, and is peculiarly in the present age, one of the primary sources of progress."

Thanks.

Questions

QUESTION: Susan Gitelson.

That was fascinating, and your presentation is beyond belief. But you left us with a dilemma, which is: What can be done about this? How do we change it? Social media is getting stronger and stronger. What about those of us who come and are only too glad to meet new people around the table or whatever and have discussion?

Cass R. Sunstein: Escaping the "Ridiculous" Social Media Bubble

CASS SUNSTEIN: The book is a pretty dark book, but there are a few ideas. There are institutions and there are individuals.

In terms of institutions, if you look at the [remarks](#) from Facebook's leadership in the last months, they are on this problem, and they recently made an [announcement](#) with respect to both "[fake news](#)" and News Feed that are alert to the advantages of an architecture of serendipity. They recently tweaked their News Feed in such a way that the related articles you're going to get aren't basically "the Daily You." It is going to be stuff on the topic, but it's not going to just tell you what you already think. It's not going to be the Colorado experiment.

The New York Times has recently created a [feature](#) which is basically "see the other side," where you have stories from different points of view, urging readers to see lots of different stuff.

Then, at the intersection of technology and democracy, there has been a flood of innovation just in maybe the last eight weeks, where you can get on your cellphone something called "Read the Other Side," which shows you as you use the app whether you're getting bluer and bluer and bluer or redder and redder and redder, and if you get really blue or really red there will come a warning saying—and don't ask me how I know that. [Laughter]

The range of entities that are doing these things is very large. So if one is involved in communications provision, as a number of people in this room are, at either a small little start-up or with something large, you can do stuff like this.

At the individual level, a lot depends on what kind of culture we have. If we have a culture that celebrates self-sorting, either through voluntary choice or through acceptance of algorithms into stuff, then to be alert that you might be actually getting into a DDD panel can shift you a bit.

Maybe in the recent election the fact that many Americans are uncomprehending of how anyone could have voted for Secretary Clinton and many Americans are, "How could someone who is sane vote for Donald Trump?"—that's a problem because the millions of people who voted for each of them, most of them don't have a problem with mental illness.

Maybe I can get a little more specific by saying that the data suggests that the percentage of Democrats and Republicans who would be unhappy if their child married someone of a different political party is actually quite high, about a third of Democrats and nearly half of Republicans, whereas a number of decades ago the percentage was trivial. In fact, people would be much more content if their child married someone of a different skin color than if their child married someone of a different political party. On one dimension that's fantastic progress, but it suggests something into which the phenomenon I'm discussing feeds. Each of us can think, That's ridiculous and we can counteract it in a little way and, since every big macro thing has micro foundations, a little act of "that's ridiculous" can move a bit the macro.

QUESTION: Hi, I'm Dick Reisman. I'm involved in media technology. I went through your new book.

One thing that I was wondering about is you wrote a really nice op-ed in 2012 where you talked about surprising validators as a way to cut through biased assimilation and bring in contrary views. I sent you an e-mail on how that can be applied in the technology of Facebook.

I'm curious about two things. One is just to expand on why. I think that's important. Facebook has the social graph so it can figure out who the surprising validators are for you and expose opinions that are from people you care about and respect who you might listen to. It seems like a huge opportunity.

I'm wondering, is there a reason that you didn't include it in the book? Have you found evidence that it's not relevant, or is it something that is promising? I've been talking to people, including Eli Pariser and other people working on filter bubbles. I think it's important if it seems promising, and, if it isn't, I'd like to know why.

CASS SUNSTEIN: The data is very supportive of surprising validators as an effective instrument. In government I think the two ugliest words I saw were "deliverable" and "validator." I hadn't heard either before. A deliverable is a product and a validator is someone outside of government who will say that what government is doing is good.

A surprising validator would be someone who would say—for the Trump administration, let's say, if someone who worked on Obama's health care plan said, "You know, Trump is actually right on this and his replacement plan is a good idea," that would be a surprising validator.

I found myself with a surprising validator a few weeks ago, someone who knew the new head of the Environmental Protection Agency and has worked with him. By my lights this was not the perfect choice to head the Environmental Protection Agency. I was told by someone who knows him, "He is a solid person. He is not an enemy of environmental protection. This is going to be just fine." Because it was someone I wouldn't expect—so that's the idea. All the data is supportive of the notion of the surprising validator.

The fact that it's not discussed in the book—it maybe should have been—wasn't not liking the concept; it's that maybe I didn't think as I was writing those pages that that was the idea that belonged in terms of the concerns of the book. That was probably a mistake.

Maybe another way to put it is that a particular kind of surprising validator is a "convert communicator" as it's called.

If you have someone who goes to a community of people, let's say, that is drawn to drug use and says, "I was a drug user. I was like you and it ruined my life." Or if you say, "I used to be a Democrat, but not anymore. I figured it out, that's a bad course"—President Reagan used that very effectively: "I didn't leave them, they left me." That was a memorable line. This is effective.

Whether Facebook should do it, I want to think longer. I think any private social media provider is alert to the risk of choosing sides, and that's good if you're a social media outlet.

You might think neutrally that "We are going to provide people with opposing viewpoints or serendipity" but less "We are going to try to persuade people that one thing is right" or "We are going to provide people with clarity that independent fact checkers have shown that this is not true"—and Facebook, my understanding is, has gone in that direction—that has more neutrality. If they tell you, "You tend to think X. Here's someone who is going to surprise you and show you that your view is wrong," that's a little more aggressive. It might be the thing to do but—

AUDIENCE MEMBER: It could be an opt-in.

CASS SUNSTEIN: It could be an opt-in, yes, definitely.

QUESTION: Don Simmons.

Fifty years ago we all read—many of us—The Hidden Persuaders. Decades ago legislatures were aware of nudges and discussed them in, for example, providing for same-day registration to vote. Utility theory has a long time ago identified risk aversion, but behavioral economics is now a much more interesting and widely followed new branch of knowledge. Why? Is it because of the studies on people and surveys? And, if so, are those sample sizes large enough to permit the conclusions that we draw?

CASS SUNSTEIN: It's a great question. I feel a little bit as if I've been asked "does the Second Amendment protect the individual right to bear arms"? I do have a view on that; it's a different topic. I'm happy to answer, but notice this book is not about behavioral economics. It does have something to do with social psychology.

Unless the rest of you object, I'll answer that question. Are you really bored by the idea of behavioral economics? There are a couple of questions there. One is "Why has it taken off?" and the other is "Are the sample sizes large enough?"

It has taken off, I think, above all because of three people—Daniel Kahneman, Amos Tversky, and Richard Thaler. Kahneman won the Nobel. Tversky would have won the Nobel with him. Thaler hasn't won the Nobel; he has been robbed. What they did was, with more rigor than anyone before had, systemize departures from expected utility theory.

Risk aversion actually is a mis-description of what people are like; they're loss-averse. With respect to gains, people are risk-seeking. With respect to losses, they are loss-averse. Risk-averse for losses, risk-seeking for gain. I may have that opposite. It depends on whether we're talking about moderate or low probabilities or high probabilities. This gets very specific. The assortment of very disciplined findings and their replication has been the Greatness with a capital G of Kahneman, Tversky, and Thaler.

In terms of the sample sizes, a lot of it started with surveys, but the core findings have now been replicated in the real world with very large samples. There are studies of whether teachers are responsive to a promise of a bonus at the end of the year? No. Are they responsive to a threat of a loss at the end of the year if their students' scores don't go up? Yes.

Loss-averse golfers do better putting for par than putting for a birdie, equivalent putts. If the average PGA golfer

putted as well for a birdie as for par, he'd make a million dollars more a year. These are putts of the same length. That's a crazy fact. It's that a par putt, if you miss it, you bogey—that's terrible. A birdie putt, if you miss it, you'll par—that's okay. That's loss aversion without a survey in the real world.

I've given you the teachers and the golfers, but there are a zillion. The replication in the real world is really important, but on the core findings we have that.

QUESTION: Ron Berenbeim.

I think the really big issue—and many people agree—politically and economically in the United States is inequality in all respects—income, education, and so on. The Internet has a tremendous capability to reduce that inequality by virtue of access to information. Is there any way this can be used in a positive way? So far it appears to have been largely negative.

CASS SUNSTEIN: It's a great question, and it's not a question on which I have, alas, expertise. I learned from Kahneman, Tversky, and Thaler to have evidence-based responses, and I don't have one for that, but I'll talk around it a little bit.

One view is that the inequality problem is the great, or one of the great, problems facing America today. I actually don't share that view. This isn't about evidence, it's about conviction. I think deprivation is the great problem.

If the upper 1 percent—imagine two societies. Imagine a society in which half the people earn $750,000 a year and half the people earn $5 million a year. There is a ton of inequality there, kind of huge inequality, but it's fantastic. Hooray for that.

The problem is in the United States the people who are at the bottom—and there are a lot of them, the poverty rate is high—the people who aren't poor but are near poor are struggling terribly. The unemployment rate, while a lot lower than it was, is higher than it should be.

I think Senator [Sanders](#) and Senator [Warren](#) have focused on the wrong percent. They focus on the upper 1 percent and the upper 10 percent rather than the bottom people. Deprivation—and it could be educational—might be the right focus, rather than "If people at the top are going through the roof, who does that hurt? Tax them more, sure, but fine."

You raised a great point about the Internet not seeming to help yet. Since all the evidence is that educational attainment is well correlated with escaping struggle, enlisting the Internet to do that would be a great idea, and it is happening. To do it in a more systematic way would be good, but that's a pretty banal thing I just said.

QUESTION: Hi, Ariana Lippi.

I was wondering if you found anything significant with people who base their convictions in religion as opposed to something like climate change.

CASS SUNSTEIN: The data I have involves now immigration, gun control, the Affordable Care Act, and one or two other areas in addition to climate change. What we're observing is regularly the Good News-Bad News Effect—as in the "how good-looking are you?"—so that people are more responsive to good news than to bad news, there is asymmetrical updating; and things that are like the opposite for the climate change, where people are more responsive to bad news than good news, but that is a less robust finding in the other areas.

If you're with me, the Good News-Bad News effect, asymmetrical updating, where people are more responsive to things that please them than things that displease them, that seems all over the political map.

The opposite finding, where people are more responsive to bad news than good news, is clear in the climate change area, and there is suggestive evidence of it in other areas, but it is not as powerful as what I told you about.

To give you an example, I would expect that if people who like the Affordable Care Act are told "it's working great, the number of people whose premiums are going up is much smaller than was thought," they would be much more responsive to that than if they heard "it's working terribly, the people are paying a ton more because of the Affordable Care Act." We are observing that, but it is less stark than what I told you about.

This is a long-winded way of saying that my expectation is the religion effect would be dominated by the Good News-Bad News Effect. There is an old study called "When Prophecy Fails," which is a study of religious believers of an admittedly exotic sort who thought that the world was going to come to an end. When it didn't, they said either, "Yes, we were responsible for it not coming to an end" or "God adjusted his plans" and they worked very hard to make sure that their beliefs were unaffected by this belief-destroying nonevent, the end of the world.

The speculation now is that the mechanisms being described—and I've had two, right?—the two mechanisms were the motivated reasoning, the "I don't like that and I don't believe that." The idea is the two mechanisms which are prior convictions being rationally the foundation for how people update.

If you read something saying "the Holocaust didn't happen, Hitler didn't kill a single Jew," that's not going to move you. That's rational. It's not necessarily "I don't like that and I don't believe that." But if you hear your child is actually much less competent than you think, "I don't like that and I don't believe that." Those are our two mechanisms.

For religious believers of whatever kind it should be exactly the same. If you're a committed atheist and then something happens that is very hard to explain except by reference, let's say, to Christianity, either you'll update, like my guy at the Waldorf, meaning you won't, or you'll say, "I don't like that, I don't believe that." If you're a religious believer, then the seeming contradiction of your religious conviction will be taken as displeasing and, therefore, not credible or as easily folded into the conviction. With respect to atheists and believers, that's daily, isn't it?

QUESTION: Jim Starkman.

Since all the forces that you've described have come to bear on the recent election and on future elections, what is your conclusion about the net direction of social media on likely outcomes going forward, since you have the broadest possible sample in the electorate and you have all the forces, positive and negative surprises, etc., at play? Are we doomed to a divided America, which you're sort of implying, but I'd just like you to clarify?

CASS SUNSTEIN: Remember Amos Tversky, one of the trilogy of behavioral heroes? He said he was an optimist and "it's rational to be optimistic, because if you're a pessimist you suffer twice, once when the bad thing happens and once when you think about it." I agree with Tversky, it's rational to be optimistic.

In terms of the arc of the United States, one finding I didn't highlight is from the climate change study. Remember the moderate climate change believers, the middle tertile? They were equally responsive to the good news and bad news. They were not asymmetrical updaters. That is the kind of solid middle of the United States with respect to climate.

With respect to things on which people don't have a clear antecedent conviction—they're not like my Waldorf guy— or they don't have a strong emotional predisposition, then the data matters. That is a hopeful sign, since so long as social media is providing people with a range of things that are there and so long as they aren't creating RRR or DDD panels, then people will learn things.

I'll give you another little hopeful sign, which predated social media but I think social media wouldn't have affected this. The previous generation's climate change problem really was depletion of the ozone layer. It has kind of

collapsed, been rendered invisible. But depletion of the ozone layer was thought to be—and not wrongly —potentially catastrophic, in the sense that the rise in skin cancer and, less alarming, cataracts would go through the roof as the ozone layer got depleted.

The idea among conservatives for a long time was that this was a crazy "fake news" thing, that Al Gore was "Ozone Man" for being concerned about this nonproblem. "How could it be that people would deplete the ozone layer? What could be crazier?"

The president who actually licked the problem and made it a non-problem is Ronald Reagan, who led the world toward acceptance of the Montreal Protocol; and that happened not with social media but with a pretty disaggregated media environment in which the depletion of the ozone layer was a laugh line. Reagan went there because of the evidence. People went into the Oval Office and said, "Here's what's going to happen if we don't do anything; here's what's going to happen if we do do something. It's going to cost money, but not huge," and Reagan said, "We've got to do it."

That is replicated in small ways in every mayor's office and in every state government, notwithstanding who you follow on Twitter.

JOANNE MYERS: I have to thank you for not only being endearing but educational and entertaining. Thank you, Cass. It was wonderful.

His book is available for you to purchase. Thank you again.