

Mental Manipulation

Thanks to the five-year age gap between our youngest and oldest daughters, my husband and I have been witnessing anew the joys of discovering a world beyond one's own mind. When our youngest daughter, Alectra, first pointed her chubby index finger at one of her unicorn figurines and asked me, "Is that *your* favorite, too, Mommy?" I smiled as I showed her the pink one I preferred, and silently cheered that she had reached a new milestone, the recognition that her own preferences may be different from mine. As she started to distinguish between happy and sad faces in her illustrated *My First 100 Words* book, I celebrated her newfound power of mind reading.

No, she isn't a telepath, as far as I'm aware. Alectra was just making inferences about the mental states of others. This important stage in cognitive development is when we become aware that others have different beliefs, intuitions, plans, desires, and intentions than our own. The ability to infer what other people are thinking and to predict what they will do as a result of their internal mental states is known as theory of mind in psychology.¹ It's called a "theory" because it's exactly that. Without actual or direct knowledge about what another person is thinking, we make inferences.² By observing the nonverbal clues of their biases, beliefs, and desires, we can not only develop a deeper understanding of their thoughts than what they may choose to share with us verbally but also predict their likely behaviors. Once Alectra discovered that I am afraid of wasps, for example, she could predict that I will dart away whenever one flies near me.³

Equipped with a theory of mind, we try to influence and persuade people around us,⁴ and use the same tools to infer when others are trying to persuade us. Before a child develops a theory of mind, they can't recognize when someone is

trying to influence them, such as when an advertiser tries to hook them on a new toy. But when they are as young as three years old, children begin to show early signs of understanding other people's intentions, and by age five they can usually understand their emotions and desires. By late preschool, they can even start to recognize embedded mental states: *Mommy is scared of wasps* (a first-order unidimensional desire), or *Daddy thinks that Mommy is scared of wasps* (a second-order embedded mental state).⁵

Children test out their newfound abilities at home. Alectra quickly pegged me as more likely to bend to her will than her father, so when she recently wanted a popsicle before dinner, she came to me with her request. "Mommy," she asked innocently, "would *I* like a popsicle?"

"I don't know, Alectra, *would* you like a popsicle?" I asked.

"Oh yes, thank you, Mommy!" she exclaimed as she reached for the freezer door. Somehow, I'd been hoodwinked again.

While everyday mind reading and persuasion are fundamental to what it means to be human, **most people feel differently about being manipulated or coerced**. We would all agree that it is wrong to hold a gun to someone's head to make them do or believe something. Or that Lord Voldemort's use of legilimency to penetrate the minds of his victims and bend them to his will in J. K. Rowling's Harry Potter series was evil.⁶ But how do we know when we have left ordinary human persuasion behind and crossed the line into the morally fraught territory of coercion or manipulation? Do advances based on neurotechnology blur that line?

Marketing to Our Brains

Are you one of the millions of people who watched the epic science fiction film *Avatar* in 2009? I was utterly transfixed by the fictional world of Pandora—the moon that humans were colonizing, threatening its indigenous Na'vi. Its premise, that a human could remotely operate a genetically engineered Na'vi body with his brain to infiltrate their community and gain their trust caught me hook, line, and sinker.

Years later, I discovered that *Avatar*'s director, James Cameron, might have hooked me in other ways, too, by using what was then the cutting-edge technique of "neuromarketing"—a hybrid between neuroscience and traditional consumer research. When Cameron was filming *Avatar*, he told *Variety* magazine that he believed its 3D aspects would be particularly appealing to

audiences because a “functional-MRI study of brain activity would show that more neurons are actively engaged in processing a 3-D movie than the same film seen in 2-D.”⁷ That spurred the firm MindSign to reach out to Cameron and offer him a freebie on their neuromarketing services. Cameron agreed, and MindSign took fMRI brain scans of research participants while they watched potential *Avatar* trailers to pinpoint which ones—and which parts of them—most engaged viewers.⁸

Moran Cerf, a neuroscientist and business professor at Northwestern University (and a leading researcher in neuromarketing for cinema) explains that good movie trailers “make all brains look alike. A good filmmaker can create an experience that will take over your brain. It is so powerful that it is actually working the same way on every other person that watches the movie.”⁹ A good film should make everyone’s brain watching it show the same patterns of attention, immersion, and engagement. Who knows if it was the stunning 3D graphics, the acting, directing, the neuromarketing, or some other qualia that is as yet unknown, but *Avatar* was nominated for nine Academy Awards, including Best Picture and Best Director, won Best Art Direction, Best Cinematography, and Best Visual Effects, and became the highest-grossing movie of all time.¹⁰

Since then, neuromarketing has gone mainstream. Neuromarketing means using physiological and brain measurements to understand consumer motivations, preferences, and decision-making to inform marketing, pricing, and product development decisions. Companies leverage the tools of neuromarketing to make their branding and products more compelling, manipulating our unconscious minds to motivate our purchasing decisions.¹¹ The field started to gain steam in the early 2000s, when researchers began to publish novel insights about consumer preferences using its tools.

Hundreds of billions of dollars are spent annually on advertising.¹² But the effectiveness of those investments has long been stymied by people’s lack of conscious awareness of the emotional, attentional and sensorial processes that underlie their choices, meaning they cannot reliably report on their reactions to products, advertisements, and brands.¹³ Brain data promises to reduce the uncertainty of traditional marketing surveys by directly decoding consumers’ unconscious preferences and biases.¹⁴ Once just a pipe dream, neuromarketing has shown the power to predict consumer behavior in a series of studies.¹⁵

One of the first groundbreaking results came in 2004, when researchers tried to understand why some people have a strict brand allegiance to Coke or Pepsi, despite their near-identical chemical compositions. In blind taste tests, people

struggle to recognize their “preferred” brand. But in unblinded taste tests, they consistently rate the brand they claim to prefer more highly. Rather than directly asking participants why, these researchers turned to their brains for answers. Participants were served Coca-Cola and Pepsi under blinded and unblinded conditions, while their brains were scanned using fMRI.

In blind taste tests, when participants drank Coca-Cola or Pepsi, they had greater brain activity in a region of the brain called the ventromedial prefrontal cortex, believed to integrate sensory, affective, and memory-related information. In the unblinded tests, people with a stated preference for Coca-Cola showed a sharp increase in brain activity in the brain regions associated with memory and cognitive control—their dorsolateral prefrontal cortex, hippocampus, and midbrain. Pepsi drinkers’ brain activity looked the same in both blinded and unblinded tests. The difference between the two groups revealed something important—Coca-Cola’s branding made those with a stated preference for Coke think and remember, and not just taste and feel when they drank a branded Coke. Pepsi drinkers weren’t having the same experience. Coca-Cola (but not Pepsi) had managed to shape user preferences in a way unrelated to the taste and smell of the beverage. This “brand effect”—an association between the brand and how it made people think and remember—changed their subjective experience of pleasure and the decision-making that followed.¹⁶

A few years later, researchers at the California Institute of Technology found a similar effect in a study involving different priced wines. The researchers served research participants wine while scanning their brains with fMRI and told them (made-up) prices for each variety. Unbeknownst to the participants, all the wines had identical retail values. The participants nevertheless consistently preferred the more “expensive wines.” Their brain activity showed a similar effect to the Coca-Cola branding. Changing the price didn’t change their sensory experience of tasting the wine, but it did increase brain activity relevant to their subjective experience of pleasure during experiential tasks.¹⁷ Armed with this information, wine producers can manipulate the price of wine without changing how it tastes and expect consumers to rate it more highly.

With insights like these, companies started to see neuromarketing as the “holy grail” of marketing—a reliable way to decode consumers’ once-inaccessible unconscious reactions to packaging and pricing and measure their actual emotional engagements with advertisements. Paying attention to an advertisement meant it had cut through other distractions. Some studies even claimed that certain brain wave patterns (such as prefrontal asymmetry in the

gamma frequency band, which you can detect with an EEG device) may be tightly correlated to our “willingness to pay” for a particular product.¹⁸

Not everyone was convinced. In 2017, Ming Su, a marketing professor at UC Berkeley, wrote in the *California Management Review* that “neuroscience either tells me what I already know, or it tells me something new that I don’t care about.”¹⁹ Most businesses have long understood that higher-priced wines are perceived as more valuable by consumers, even when they taste the same. Su was far from the only dissenting voice. Academics were generally pessimistic in the early 2000s. Their pessimism was buttressed by early setbacks in the field, including inappropriate assumptions about the localization of particular psychological processes within the brain, and a lot of oversimplification about how we make decisions. The technology and software powering it also limited neuromarketers’ ability to measure brain activity and analyze it accurately, and the field was stymied by methodological problems, including reverse inference, whereby researchers assumed a pattern of activation meant a certain cognitive process was engaged.²⁰

But as the science, technology, and methodology supporting neuromarketing have progressed, brain data now offers real advantages when predicting the future success of consumer products. More portable fNIRS are being used to map areas of the brain engaged with ongoing psychological processes. EEG has become a popular mainstay in neuromarketing because of its ability to detect unconscious preferences and biases, despite some of its limitations, including its inability to pinpoint where in the brain activity occurs (including its inability to reach the deep, subcortical regions of the brain, where a lot of consumer decision-making occurs).²¹ Physiological measures—heart rate, electrodermal activity, eye tracking, facial electromyography, and the like—complement these tools, enabling marketers to detect preferences, desires, and biases outside of individuals’ awareness and control.²² By 2017, the Advertising Research Foundation, a nonprofit industry association for creating and sharing knowledge in advertising, heralded neuroscience as better able to predict what consumers actually want and do.²³

Today there is a burgeoning industry of more than 150 neuromarketing firms worldwide, ranging from technology providers to those that offer more comprehensive services.²⁴ Some companies, like NBC and Warner Bros. Discovery, have operated their own neuromarketing units for years, while others, like Microsoft, Google, and Meta, have formed units to do the same. Most of these companies believe that decoding the subconscious consumer mind is

critical to their ongoing success.²⁵ It seems to be working.

Anyone who watches TV sees a lot of PSAs promoting worthy causes. You may even have responded to some of them with donations, to help families affected by wildfires, or those impacted by the tragic loss of life in a mass shooting in the United States. You may have donated to support Ukrainian refugees, what with more than two-thirds of the country's children displaced by the Russian invasion.²⁶ The United Nations has called it the "fastest and largest displacement of people in Europe since World War II."²⁷ But maybe not, if the call to action included images of their displacement's devastating effects. Why is that?

The *New York Times* reporter Charles Duhigg puzzled over a similar issue—why people weren't donating to Syrian refugee relief. One answer came from his interviews with the social scientists Jennifer van Heerde-Hudson and David Hudson, who have spent years studying how charities solicit donations. "Children who have lost their homes, starving families, the heartstring things," David Hudson told him. "That's what everyone believes works." But they found the opposite to be true. When campaigns shift from images of poverty-stricken children and messages like "Please donate before it's too late" to hopeful and inspiring images of children holding signs like FUTURE DOCTOR, people are more likely to give. "If you can trigger a sense of hope, donations go up," explained Mr. Hudson.²⁸ Or as Duhigg puts it, "It's not entirely your fault" if you aren't donating to refugees. "You just haven't been manipulated properly."

When neuromarketers tweaked an unsuccessful campaign by the Italian UNCHR for refugees, its new commercial led to a 237 percent increase in sellable calls over the prior one. The brains of test subjects showed them how to do it. The first commercial had low emotional arousal throughout, and poor engagement during the final call to action. Using EEG insights from participants watching the commercial, they modified the new commercial with new images to evoke greater empathy in viewers, and with new visual effects in the call to action that better engaged viewers' brains.²⁹

Is Duhigg right to call this "manipulation"? The Advertising Research Foundation was concerned enough that it called for government standards to guide neuromarketing research.³⁰ The Neuromarketing Science and Business Association—established in 2012 to support researchers and practitioners of neuromarketing worldwide—developed a code of ethics to guide responsible progress in the field.³¹ Neuroethicists have long called for protections for people who might be harmed or exploited by neuromarketing as its effectiveness

improves, potentially eroding viewers' autonomy.³² But there is little consensus on what constitutes permissible and impermissible uses of neuromarketing, despite growing anxiety about whether researchers might discover aspects of the human brain that could “turn individuals into buying robots and produce dangerous behaviors,” like addiction or overconsumption.³³

As giant corporations like Coca-Cola, McDonald's, and Procter & Gamble seek to understand consumption patterns by vulnerable populations such as children, addicts, and gamblers, many worry this research will be used to harm humanity. Anxiety over neuromarketing has been further amplified by its opaque use by tech giants to shape customers' experiences. Meta has already come under fire for running at least one psychological experiment without user consent—manipulating the moods of more than seven hundred thousand users by altering its newsfeeds with happy and sad content, just to see the results.³⁴

Can a trailer, movie, or product be so enticing that we can't reasonably resist it? Does it matter if the tactic is put to uses that are intended to benefit versus harm us? Does it matter if the influence is hidden from our view?

Persuading or Addicting the Brain?

In 2017, the entrepreneur, venture capitalist, and former Facebook engineer Justin Rosenstein joined several other former Facebook executives to sound an alarm about the techniques social media companies deploy to target users' unconscious decision-making.³⁵ Rosenstein helped create the Like button (originally called the Awesome button) during his days at Facebook, which later became a standard feature across most social media platforms. It was intended, he explained, to “send little bits of positivity” across the platform.³⁶ But Rosenstein had since come to believe that it was profoundly harming humanity by addicting people to the platform and tying their self-esteem to the Likes they receive.

Rosenstein has joined the ranks of other leading technologists who are weaning themselves off products they helped to create, while sending their children to schools where iPhone, iPads, and even laptops are prohibited. I can't blame them. There is something profoundly disconcerting about having your barely verbal two-year-old wake up and scream, “I need your phone!” as her first words of the morning.

By 2018, the average Generation Z smartphone user was so enamored with their phones that they unlocked them at least 79 times a day.³⁷ Most of us are

using our phones at least 20 percent more often than we did in 2015,³⁸ while 60 percent of college students believe they have become an addiction, and 87 percent of millennials admit that their smartphones never leave their side.³⁹

There is growing concern about the societal implications of this addiction, not the least of which is the interference with our ability to focus.⁴⁰ “Everyone is distracted,” Rosenstein said. “All of the time.”⁴¹ Emory law professor Matthew Lawrence recently took on these issues by arguing for a right to freedom from addiction, citing the dangers of social media addiction and concern that “social media and game companies have knowingly used technology to plant repetitive, unwanted thoughts in users’ minds, without their knowledge and consent; indeed, without even the basic ‘warning: this product is addictive’ that now appears on cigarette packages.”⁴² With harms ranging from increased risks of suicide to automobile accidents from distracted drivers and loss of workplace productivity, Lawrence argues that the US Constitution should be interpreted to address “the liberty implications of addictive technology” to bolster the case for legal interventions.⁴³

But just as technologists like Rosenstein are abandoning ship, other technologists are paying high-ticket entry fees to attend conferences curated by Nir Eyal, the author of *Hooked: How to Build Habit-Forming Products*, to learn how to better addict people to their products.⁴⁴

The everyday technologies we use “have turned into compulsions, if not full-fledged addictions,” Eyal writes. This is far from accidental, but “just as their designers intended.”⁴⁵ His website advertises “how to build habit-forming products” and invites companies to “Discover the secrets the world’s leading technology companies use to keep users coming back—and apply them *right now* to your product.”⁴⁶ Facing public scrutiny, by 2017 he started to acknowledge the growing anxiety about whether these were manipulative tactics, cautioning his audiences that they should be careful not to use them for harm. But he still stalwartly defends what he teaches, arguing that just as “we shouldn’t blame the baker for making such delicious treats, we can’t blame tech makers for making their products so good we want to use them.”⁴⁷

Tristan Harris, a former Google employee turned tech critic, cautions that we are “jacked into this system ... All of our minds can be hijacked. Our choices are not as free as we think they are.”⁴⁸ Like Eyal, Harris studied under B. J. Fogg, a behavioral psychologist at Stanford University who is well known for his mastery of the ways that technological design can be used to persuade. Their similarities end there. Instead of teaching Google how to exploit our brains,

Harris cautioned them against doing so. As a result, he became their in-house design ethicist and product philosopher. I “got to sit in a corner and think and read and understand,” he says. That’s when he came to believe that LinkedIn exploits our need for social reciprocity to widen its network, how media platforms like Netflix use autoplay to keep our attention fixed as the service transitions from one episode to the next, and how Snapchat created Snapstreaks to keep communication between at near constant levels between its users.⁴⁹

Each of these approaches exploits shortcuts in our brains. To keep ourselves safe, we pay more attention to fearful, dangerous stimuli,⁵⁰ so social media uses notifications and alerts to make our brains believe that we urgently need to turn our attention back to their platforms.⁵¹ Children with a particular genotype are more likely to be addicted to nicotine if they vape their first ecigarette before they are fifteen years old;⁵² it’s perhaps knowledge of this which led tobacco companies to use flavored tobacco products and advertisements to target young people to addict them for life. Features like infinite scroll and algorithmic recommendations prey on evolved mechanisms in our unconscious without us even realizing they are doing so.⁵³ The algorithms have become so sophisticated that tech companies can even target specific people to hook them on their platforms.⁵⁴ “Once you know how to push people’s buttons, you can play them like a piano,” Harris declares.⁵⁵

Companies have been doing some variation of this all along. Food companies have long manipulated combinations of salt, sugar, and fat to serve us foods that our bodies end up craving.⁵⁶ Politicians have long tailored their manner of dress and speech to make us believe they are people like us and hence trustworthy, to garner our votes at the polls. I remember a course I took to prepare for the MCAT, where they counseled us to write in cursive rather than print for the essay portion of the exam, because cursive writers statistically scored higher. Have persuasion technologies just gotten better at giving us more of what our brains have wanted all along?

How, if at all, is this kind of persuasion different from Alectra turning the tables on me to make it seem like it was *my* idea to offer her a popsicle before dinner? Why do we smile when a young child gets her way but shudder when tech giants do the same? Is it because the child’s manipulation is transparent while theirs is harder to detect and resist?⁵⁷

Taking Advantage of Brain Heuristics

For years, I have studied when and why criminal defendants use brain scans or neuroscientific experts to argue that their “brain made them” commit a criminal offense. If you believe that all our actions, perceptions, and beliefs ultimately come from our brains, then describing a defendant’s brain in greater details shouldn’t have that much bearing on their culpability. But defenses that describe criminal conduct as arising from the brain rather than a bad choice regularly throw judges and juries for a loop.

Some of our credulity may be explained by what has been called the “seductive allure of neuroscience,”⁵⁸ a phenomenon in which people are swayed to think more favorably about psychological explanations if they include references to or images of the brain.⁵⁹ Even if the references are logically irrelevant or make a good explanation worse, people still rate arguments that include neuroscience as more persuasive. Do people just favor longer over shorter explanations? Or explanations that seem more authoritative? Perhaps scientific-sounding jargon is more inherently persuasive?

Dr. Deena Weisberg, a psychologist at the University of Pennsylvania, set out to find out. In a series of three experiments, she tested each of these hypotheses. She recruited participants to take online surveys in which they were shown good or bad versions of explanations. Each example included a description of a psychological phenomenon (such as that babies have the ability to do simple math; that there are gender differences in spatial reasoning; and that there are differences in seeing and imagining objects) and gave eight different explanations for each. The explanations were presented in two lengths. Some included neuroscience, and some made no reference to science at all. Each trial participant saw one of the phenomena, received one explanation, and was asked to rate that explanation on a seven-point scale ranging from very unsatisfying (−3) to very satisfying (+3). Subjects rated longer explanations as better than shorter ones, but neuroscience had an independent and stronger effect. While people could judge between good and bad explanations, adding neuroscience interfered with their ability to do so.⁶⁰

Researchers have replicated these findings across other reducible scientific disciplines and found similar effects on our ability to tell a good argument from a bad one. To understand the significance of reducibility, “consider the relationship between chemistry and physics,” explained Weisberg. “While it is logically possible that atoms (and other elements of a physical ontology) could exist without there being any molecules, it is not logically possible that molecules could exist without atoms. Atoms, then, are logically prior to

molecules. If an explanation of a molecular phenomenon is translated in terms of atoms, and if the atomic translation does not omit any aspect of the molecular version of the explanation, then we can say that the explanation has been *reduced* from the chemical to the physical level.”⁶¹ When we are presented arguments with reductive reasoning, even incorrect reasoning, it makes it harder for our brains to pick out a good argument from a bad one.

While no one has yet established exactly why that may be, reductive explanations may misguide us into failing to think—using shortcuts our brains provide us instead, that make us more likely to believe the information before us without thinking critically about it.⁶²

We use all kinds of cognitive shortcuts to help us navigate our environments more efficiently, and reductive scientific logic appeals to a hardwired heuristic that allows us to categorize reductive-sounding logic as correct.⁶³ Reductive logic may lead us to accept an argument quickly rather than carefully, and “people who believe false things ... don’t think carefully,” claims David Rand, professor of Management Science and Brain and Cognitive Sciences at MIT.⁶⁴

Using reductive science and other cognitive shortcuts as heuristics for judgments about argument quality makes us more susceptible to fake news and other misinformation. People spread misinformation intentionally and unintentionally. Those who do so intentionally often study how to frame their claims in ways that capture our attention and make us more likely to pass that information on to others—with sometimes devastating effects.⁶⁵ Misinformation has tanked stock prices, wiping out, in one case, \$130 billion in stock value after a false tweet about Barack Obama being injured in a White House explosion went viral in 2013.⁶⁶ Misinformation about terrorist attacks or natural disasters can set off panics, undermining entire societies.⁶⁷

Splashy headlines are another strategy that appeals to our unconscious cognitive biases. The more outlandish the claim, the more likely our brains are to focus on the novel stimulus.⁶⁸ Unexpected information can pierce our attentional filters by triggering our sensory cortex to pay attention, triggering the brain to release dopamine, which makes the attention that much more rewarding.⁶⁹

These strategies to exploit our unconscious biases may partly explain how false information spread so widely during the COVID-19 pandemic. At the height of the pandemic, the not-for-profit Center for Countering Digital Hate found that anti-vaccine activists had reached more than fifty-nine million followers with messages that were shared over 812,000 times on platforms including Facebook, YouTube, Instagram, and Twitter. More than 65 percent of

that anti-vaccine content was attributable to a group they dubbed “the Disinformation Dozen,”⁷⁰ many of whom have been spreading spurious medical claims for years.⁷¹ Some of the most prominent members used A/B testing to see which of their stories were most likely to go viral. If we were to study the most successful of those claims, they likely preyed on unconscious neural processes, causing real and devastating harm to individuals and society. Exposure to even a small amount of misinformation reduced the number of people willing to take a COVID vaccine by up to 8.8 percent.⁷²

We might believe we are impervious to tactics designed to make us susceptible to misinformation. But even the most well informed of us fall prey to them. The more we are exposed to the same false claims, the more our brains begin to mislead us into believing them. Becoming more deliberate and critical about what we read and hear can help us counteract those effects, making us less likely to contribute to the problem.⁷³

What Mental Manipulation Is Impermissible—and When?

We are constantly bending and being bent to the will of others. And neurotechnology may be enabling newfound ways for those seeking to bend others to their will.⁷⁴ In chapter 3, I discussed the report by Dr. Ahmed Shaheed, the special rapporteur to the UN General Assembly on freedom of religion or belief, and his recommendation to expand the international human right of freedom of thought to include the right not to reveal one’s thoughts nor to be penalized for them. He also recommended that freedom of thought include the right not to have our thoughts manipulated.⁷⁵ But manipulation is a slippery concept. If ill-defined, an absolute prohibition on it could do more harm to human interactions than good.

About a decade ago, I went down a deep rabbit hole when I was trying to untangle claims about philosophical and legal free will. The written debate goes back at least two thousand years, but neuroscientists have recently joined the fray by arguing that our decision-making is hardwired in our brains.⁷⁶ Punishment, they argue, cannot be justified by retributivism—an eye for an eye—because people are not morally culpable for their actions. I disagree and have sought in my own scholarship to explain why freedom of action is a freedom worth defending.⁷⁷

In a well-known 1971 essay titled “Freedom of Will and the Concept of a Person,” the American philosopher Harry Frankfurt describes what he calls a

peculiar characteristic of humans—that we can form “second-order desires.” Besides our subconscious preferences, biases, and desires, we can also “want to have (or not to have) certain desires and motives.”⁷⁸ Frankfurt calls this capacity for reflective self-evaluation of those biases and desires higher-order volition. We don’t have to be fully *aware* of our unconscious desires to engage in reflective self-evaluation. We might be completely unaware of some desires, while being mistaken about others. Free will, he argues, is our capacity to form higher-order volitions, by recognizing certain desires as our own.

Frankfurt uses an example of two animals addicted to drugs. One is conflicted about his addiction—he craves the drug but also wants to be free from it. He wants his desire to be free from his addiction to become the one that drives his behavior. The other animal also has conflicting desires but lacks the capacity for self-reflection, and so doesn’t form a preference between them. The first animal is human while the latter is not, because only the first makes one of his desires “more truly his own, and in so doing, he withdraws himself from the other.” Frankfurt implicitly connects this to manipulation, by explaining that when the human addict is unable to break his addiction, he feels like the force “moving him to take the drug is a force other than his own.”⁷⁹ When we believe that something *other* than our free will is driving us to act contrary to a desire we identify with, we feel as if we are being manipulated.

Frankfurt’s example helps us distinguish between freedom of will and freedom of action. Freedom of will is our capacity to identify with our desires. Freedom of action enables us to make our will our own through our actions. Our freedom of will may be illusory—we commit to desires, biases, or preferences believing we have done so freely, but we may have chosen that preference because it was unconsciously primed by our environment. Our freedom may also be interfered with, making it harder to make our volition effective, if we are manipulated into acting compulsively with a “force other than [our] own.”⁸⁰ We may want to stop checking Instagram every five minutes, but cleverly timed notifications compulsively draw us back in.

In *Autonomy and Behavior Control*, Gerald Dworkin characterized a person’s motivation as belonging to a person without it truly being “their” motivation, if that motivation is brought about by interfering with their ability to reflect rationally on their interests through deception, or by short-circuiting their desires and beliefs, making them a passive recipient of the change.⁸¹ Philosophers Daniel Susser, Beate Roessler, and Helen Nissenbaum in a recent article defined manipulation in the digital age,⁸² arguing that permissible influence appeals to

our “capacity for conscious deliberation and choice,” while manipulation takes “hold of the controls,” depriving us of “authorship over [our] actions” and driving us “toward the manipulator’s ends.”⁸³

Other scholars describe manipulation as interfering with our “mental integrity,” which Andrea Lavazza describes as “the individual’s mastery of his mental states and his brain data.” He argues that we should draw a bright line that prohibits unconsented-to interferences that “can read, spread, or alter such states and data in order to condition the individual in any way.”⁸⁴ Marcello Ienca and Roberto Adorno are more tempered in their claims, arguing for “specific normative protection from potential neurotechnology-enabled interventions involving the unauthorized alteration of a person’s neural computation and potentially resulting in direct harm to the victim.”⁸⁵

These accounts all coalesce around a definition of manipulation as hidden attempts to use our cognitive biases, emotions, or subconscious “as vulnerabilities to exploit” by bypassing our capacity for conscious thought.⁸⁶ What they get wrong is that they build on an outdated Freudian view that our psyche has “two minds”—a conscious and an unconscious one. We have since learned that unconscious processes use the same brain regions in the same ways as conscious processes. Our unconscious mind is primed *all the time* through regular “strength” stimuli (rather than hidden and subliminal ones). Think of the popcorn and soda advertisements before a movie begins. They are hardly hidden, but they play to our baked-in desires. Advertisers and tech giants have just gotten much better at identifying and targeting them. Indeed, social psychologists have argued for decades that people are unaware of the powerful influences that are brought to bear on their choices and behavior.⁸⁷

Tanya Chartrand, a Duke professor, provides a different way to understand the problem by distinguishing between different kinds of unawareness—of primes that triggers our mental process, of the mental process itself, or of the consequences and effects of those triggers. Yale professor of psychology and cognitive science, and professor of management John Bargh, who has long studied unconscious influences on the consumer mind, argues that it’s the last of these we focus on—“not whether the event itself is perceived consciously or not (it almost always is), [but] whether the person is aware of how that event *affects* their choices and behavior.”⁸⁸ Poets, politicians, governments, and advertisers all know how to exploit the unconscious influences on our lives, he explains. And advertisers know well how to manipulate hidden mechanisms that drive our behavior. But people are often unaware of how those influences impact their

actions. Which is why it's critical that we understand what others can and can't do to change our minds.

In chapter 8, we'll consider the starkest examples of manipulation perpetrated by assaulting our brains, which clearly violate our right to self-determination and freedom of thought. The more difficult cases to resolve, however, are the subtler influences that shape our everyday decision-making and that are quickly becoming normalized.

It's much easier to prime us to act in ways that are consistent with our existing goals. Advertising a weight-loss program to a person who is trying to gain weight will inevitably fail. Priming us with cues that are related to our goals, however, will focus our "selective attention" on "goal-relevant features of the environment," which can shape our choices that follow.⁸⁹ Professors of marketing and psychology Gráinne Fitzsimons, Tanya Chartrand, and Gavan Fitzsimons found compelling evidence of this effect when they subliminally primed study participants with Apple and IBM brand logos. The Apple logo prime led people to act more creatively on subsequent study tasks compared to subliminal IBM logo priming—but only when creativity was a part of the participants' self-descriptions.⁹⁰ Apple evoked in these participants an association of creativity, leading those with a prior stated goal of being creative to act more creatively on subsequent tasks. Because IBM didn't evoke the same association, even those with creativity as a stated goal didn't act more creatively when primed with IBM instead.

Subliminal priming can also nudge us to act in ways that are irrelevant to goal-directed behavior. Bargh describes a study in which participants had been unknowingly primed with words related to betting (bet, gamble, wager) or standing pat (pass, fold stay) on given hands in games of online blackjack. The participants' betting was statistically consistent with the primes they had been given. But they believed that they had freely and consciously chosen their bets—and the belief was even stronger when they were primed than when they were unprimed.⁹¹

Even asking us questions about our hidden vices can change our subsequent behavior.⁹² We often have conflicting attitudes about behaviors like smoking, drinking, and using drugs. We get a short-term reward (like a dopamine hit in our brain) when we indulge, but we also understand the negative long-term consequences that go with them. When we hold conflicting explicit negative and implicit positive attitudes about a behavior, priming may give us "license to sin."⁹³ Frankfurt's human addict wants to break his addiction but asking him how

often he plans to take the drug in the next week can nudge him toward doing so more often, despite his explicit preference otherwise.⁹⁴ When researchers asked students about their attitudes toward skipping class, they reported strongly negative attitudes toward doing so, but then skipped class more frequently in the weeks following.⁹⁵ When study participants were asked how often they would go out drinking or watch television instead of studying, they also did so more frequently in the week following.⁹⁶ But when framed negatively—telling participants that drinking and wasting time watching television are vices to be avoided—the vice behavior remained the same.⁹⁷ How an influencer frames a question can liberate us to sin or increase our ability to avoid doing so.

All of which makes it exceptionally unrealistic at best, or outdated at worst, to define unlawful manipulation as intentionally using hidden influences to affect our decision-making. When neuromarketers use advances in neurotechnology to discover what makes us tick, and then use that information to make their products more enticing, they don't render us unable to act consistently with our goals. As of yet, no one has discovered the so-called buy button in our brains. When the Disinformation Dozen exploit evolutionary shortcuts in our brains to make us more susceptible to fake news, they don't prevent us from getting vaccinated, even if their bad arguments do appeal to our heuristics.

But when Nir Eyal teaches companies how to addict us to their products, despite the long-term negative consequences of doing so, we should make sure we retain the ability to act otherwise and investigate whether they intend or cause actual harm as a result. If a product becomes actually or nearly impossible to resist, our freedom of action will be hindered and our self-determination and freedom of thought will be put at risk.

Dr. Shaheed concedes that freedom of thought cannot and should not be used to prevent “ordinary social influences, such as persuasion.” We may encourage others, advise them, even cajole them, he argues. But at some point, an influence crosses the line from permissible persuasion to impermissible manipulation.⁹⁸ He offers a nonexclusive set of factors to consider, including whether the person has (1) consented to the practice with fully and freely informed consent; (2) if a reasonable person would be aware of the intended influence; (3) if there is a power imbalance between the influencer and target; and (4) if there was actual harm to the person subject to manipulation.⁹⁹

These are helpful but still don't make clear the *nature* of the influence we are defending ourselves against. We can't and shouldn't attempt to regulate every

marketer, politician, artist, or entity who tries to appeal to our unconscious biases, desires, and neural shortcuts, lest we interfere with everyday interactions that are part of what it means to be human, whether those attempts are hidden or visible, or targeted at our unconscious or conscious neural processes. But when a person or entity tries to override our will by making it exceedingly difficult to act consistently with our desires, and they act with the intention to cause actual harm, they violate our freedom of action, and our right to cognitive liberty should be invoked as a reason to regulate their conduct.

However begrudgingly, we must admit that neuromarketing per se does not violate cognitive liberty, so long as the research is conducted ethically and the findings are not used to intentionally cause us harm. We can't say the same about intentional efforts to exploit our brains by addicting us to technology, social media platforms, or other products. While our brains may fall for bad arguments when cleverly framed, we can and should encourage societal interventions that nudge us to slow down and think critically. When Twitter asks "Would you like to read the article first?" before retweeting it, it's asking us to slow down and think critically before we act. More companies ought to implement mechanisms that encourage users to do the same. And we should aspire to do so ourselves even when we aren't nudged to do so. But freedom of thought shouldn't be used as an excuse for filtering that information for us.

As for Dr. Shaheed's recommendation that we consider whether a person has freely and voluntarily consented to an intervention? While consent will rarely be enough to shield us from the coming encroachments to cognitive liberty, with at least the newest technique we turn to next, it should be a critical factor in considering the legitimacy of the technique.

Bypassing the Conscious Brain

A few months ago, I had a fitful night of sleep plagued by intrusively vivid dreams in which I was consoling a friend. I woke up in a state of agitation. Suddenly my phone buzzed with a notification. It was that same friend—whom I had not spoken with for many months—asking if I was okay. "It's so strange to hear from you this morning," I replied. My phone dinged again. "I had an incredibly vivid and disturbing dream about you," my friend continued. "I dreamt you had died. I was attending the memorial service. I was inconsolable." I was deeply shaken. Not just by the idea that my friend had dreamed I was dead! I was much more disturbed by the seeming synchronization between our

dreams. What could that possibly mean? Was it just a random coincidence? (I rarely dream about this friend, and suspect my friend rarely dreams about me.) I then had an even more disturbing thought: What if it wasn't a coincidence at all? What if our dreams had somehow been tampered with? That may be less far-fetched than you think.

For years, Coors has been locked out of showing commercials during the NFL Super Bowl game, because the league has a contract with Anheuser-Busch. So, Coors found a new way to infiltrate the minds of NFL onlookers. With help from Dr. Deirdre Barrett, a psychologist who is an expert on dreams, it created a film with specific audio and visual stimuli. In exchange for half off a twelve-pack of Coors beer (or a free twelve-pack if shared with a friend), Coors invited participants to watch a ninety-second video featuring images of mountains and Coors beer right before going to bed, and then listen to a soundscape while they were sleeping. When they woke up, they were asked what they had dreamed about.¹⁰⁰ Dr. Barrett claimed the “participants reported similar dream experiences, including refreshing streams, mountains, waterfalls, and even Coors itself.”¹⁰¹

How could this possibly work? Upon awakening from a dream, we have a period of about twenty minutes before blood flow is fully reestablished to the dorsolateral prefrontal cortex, during which we are more suggestible. Dr. Barrett and other researchers capitalize on this to incubate dreams. Suppose you decide, for example, that you want to work out some emotional turmoil you've been experiencing while sleeping. You could record a dream prompt message for yourself using your own voice. That prompt will then be played back to you as you fall asleep. After falling asleep, a sleep sensor would wake you up and play your recording to guide your thoughts during the time your brain is in a suggestible state. You would then fall back asleep only to be awakened again after a predetermined time and prompted with the same dream message. This cycle of sleep and re-prompting would continue throughout the night to guide your dreaming.¹⁰²

Coors is not the only company that is using dreams to sell things. Xbox's “Made from Dreams” was created to promote the Xbox Series X. The videos were crafted from dreams selected streamers reported after playing the new console for the first time.¹⁰³

This seeming sci-fi scenario is made easier by sensors that can detect brain wave activity, pinpointing the stages of sleep we are in at any given moment.¹⁰⁴ Even a bedside smart speaker can be used to detect the breathing patterns

indicative of different stages of sleep, triggering it to play soundscapes that can incubate dreams.¹⁰⁵ Researchers have even communicated directly with lucid dreamers—the stage of sleep where one is aware while dreaming—and had them answer questions and solve math problems.¹⁰⁶

There is great potential to these advances—to treat nightmares, enhance learning and memory consolidation, even overcome PTSD and addiction.¹⁰⁷ But dream researchers also caution about the perils of doing so. Collectively, they have started to author ethical norms and guidelines because of “the threat of the capture, sale and colonization of the dreaming self in the form of data; the outsourcing of introspection, trusting sensors more than senses; and the infiltration of our most private spaces by those who may wish to harm or manipulate us.”¹⁰⁸

In June 2021, forty sleep and dream researchers published an open letter in which they stated that targeted dream incubation “is not some fun gimmick, but a slippery slope with real consequences.”¹⁰⁹ One of its lead authors, cognitive scientist Adam Haar, had invented a device that tracks sleep patterns and guides wearers to dream about specific subjects by playing audio cues. He was alarmed when companies ranging from big tech to major airlines started to ask him for help incubating consumers’ dreams. He worries that people are particularly vulnerable to suggestive content when sleeping and fears that the absence of regulations for in-dream advertising could lead to a future in which we “become instruments of passive, unconscious overnight advertising, with or without our permission.”¹¹⁰

But other researchers, like Tore Nielsen, a dream researcher at the University of Montreal, are far less troubled by those dystopian scenarios, because the interventions are unlikely to work unless the dreamer is aware and willing to participate in the dream incubation. “I am not overly concerned,” he said, “just as I am not concerned that people can be hypnotized against their will.”¹¹¹

I didn’t agree to dream about my friend that night. I suspect it was just a random coincidence, or that something had unconsciously primed us both to think about each other earlier that day. But the experience did remind me of the coming possibilities and reaffirm in my mind that at the very least, dream incubation is an intervention that ought to be consented to anytime it is used. It’s one thing to target our unconscious processes when our conscious processes are active and can be used to defend our freedom of thought. But it’s much more troubling to have our unconscious processes infiltrated when our conscious processes are silenced during sleep and cannot be used to filter information.

Nonconsensual dream incubation—whether used to benefit or harm us—falls much closer to the wrong end of the spectrum between innocent influence and intentionally assaulting the brain, the practice we turn to next as a stark violation of cognitive liberty.