

“Trusting Your Untrustworthy Self”

Michael Brownstein

Columbia University School of General Studies Honor Society Induction Keynote Address

Thanks so much to Dean Awn and Professor Rosner for inviting me here today. I’m so honored and grateful. More importantly, congratulations to the inductees – this is your day, and it’s well-deserved.

Now, it would be sort of crazy for me to try to give you the usual sort of advice one gives in this setting. For one, I suspect there’s not a single thing all of you have in common, except the fact that you’re kicking butt at a difficult and prestigious university. I also can’t tell you what the “real” world is going to be like, since many of you have, astoundingly, worked full time or raised kids while crushing it here at school. Nor can I tell you about receding hairlines and mortgages and boredom and loneliness and all the other challenges that come with leaving your college years behind, since probably half of you are older than me.

Instead of any of that, I’ll talk about The Lego Movie. In particular, I’ll give you an argument. The argument is that the Lego Movie is smarter than Leo Tolstoy. I hope at least you’ve heard of the Lego Movie.

Here’s something that Tolstoy wrote in his novel *Anna Karenina*:

“Reasoning led him into doubt and kept him from seeing what he should and should not do. Yet when he did not think, but lived, he constantly felt in his soul the presence of an infallible judge who decided which of two possible actions was better and which was worse; and whenever he did not act as he should, he felt it at once.”

Listen closely here, since Tolstoy is saying something strange. “*Reasoning* led him into doubt,” not cowardice or fear or ignorance or any of those things. “When he did not think, but lived,” he acted as he should. What’s going on here?

In fact, Tolstoy is anticipating an idea which is all the rage right now. In short, the idea is that your gut instincts are pretty smart. You’ll find this idea in scholarly work in psychology, in philosophy, and in popular science writing too (not to mention self-help books).

Consider something called the Iowa Gambling Task. It’s a psychological test that asks you to turn cards over from 4 face-down decks. The cards tell you whether you’ve won or lost money, sometimes real money, and your goal is to win as much money as possible. Unbeknownst to you, two of the decks are “bad”—their cards offer high initial rewards but overall losses—and two of the decks are “good,” offering lower initial rewards but overall gains. Along the way, an experimenter asks you if you like any of the decks best and, if so, why.

How many card-turns do you think it takes until most people figure out the pattern? The answer is that it takes about 80 turns until most people know which are the good decks and are able to describe the patterns of initial versus overall gains and losses. But here’s the fascinating finding: after only 20 turns, people exhibit signs that they can distinguish the good decks from the bad decks. They just don’t know that they

know it. For example, you unknowingly sweat a little more before picking from the bad deck. That is, you sweat more before picking from the deck with the high-risk cards, because you're nervous, even though you don't fully consciously realize it. Antoine Bechara—the researcher who invented the Iowa Gambling Task—takes this to show that there are times when our unconscious impulses are smarter than our reasoned choices.

As I said, this idea has gone beyond academic psychology. Science writers like Malcom Gladwell have popularized the idea of “thinking without thinking” and “the power of the unconscious.” And similar ideas have found their way into philosophical writing on ethics and virtue, the basic idea of these being that we sometimes act better—more courageously or compassionately, for example—when we act spontaneously. Imagine that you're visiting a friend in the hospital—this example comes from a philosopher named Michael Stocker—and she asks, “why did you come?” Would she prefer that you answer, “I deliberated carefully upon my moral duties and I determined that they require me to visit you in the hospital,” or would she rather you just say, “you're my friend, it just seemed right?”

Ok, so Tolstoy was a genius ahead of his time. Shocking, right? But as you might be thinking, there's surely a problem here. Our instincts can't be all good. And here's where we get back to the Lego Movie. The movie's protagonist, Emmett, needs to become a master builder—yes, a master builder of Legos—just stay with me here. The master builder's skill is almost magical. Master builders just see immediately what needs to be done as if by instinct. They're like the Daoist master carpenter who just “sees” the next cut in the wood, or like the basketball pro who passes so well it's like she has eyes in the back of her head. The master builders have, you might say, “transcended the rules.”

Emmett, however, in the Lego Movie, is only good at following the rules. He just can't get this instinct thing, no matter how deep and profound the advice he receives from the old blind wise man, Vitruvius. And so in the climactic moment—with the bad guys closing in! and the Lego space ship falling apart! and certain death immanent!—Emmett has *tried* to “trust his gut” and failed yet again, making a mess of everything, and what does Vitruvius say? You have to imagine this in the voice of Morgan Freeman:

“Emmett, trust your instincts, unless your instincts suck!”

Exactly correct, Morgan Freeman/old blind wise man Lego! You *should* trust your instincts—just like Bechara and Gladwell and many serious academic researchers suggest—unless your instincts aren't very good.

Tolstoy probably realized this too, but, well, that's inconvenient for my argument, so I'll just leave it aside.

Vitruvius' advice is unhelpful in an obvious way to Emmett, of course, since Emmett's whole problem is that his instincts really aren't very good. But it gets worse, I have to tell you. Alongside the evidence that our gut instincts can be smart, there has also been, shall we say, a rather extensive set of findings in cognitive and social psychology over the past 40 years showing us just how bad our instincts often are. Some might say that these findings have revised our picture of the mind on the

scale that Darwin revised our picture of biology and natural history. I'll give you three striking examples. You may be familiar with some of them.

Imagine Linda. Linda is 31 and she's outspoken and smart. She was a philosophy major (of course!). In college, Linda was very concerned with issues of social justice and discrimination, and she attended several anti-nuclear demonstrations. Which do you think is more probable: Linda is a bank teller, or Linda is a bank teller and a feminist? (Hands.)

Those of you who said that it's more probable that Linda is just a bank teller are right, but you probably already know this example and are a bunch of no-good cheaters. Just kidding. The rest of you, who said that Linda being a bank teller and a feminist is more probable than Linda being just a bank teller, have made a mathematical error. Don't worry though, almost everyone does. Anything *A* *has* to be more probable than *A and B*. Imagine it this way: which is more probable, that today is Tuesday or that today is Tuesday and that it's hot outside? Anytime it's Tuesday and it's hot outside, it's also just Tuesday. But sometimes it's Tuesday and it's not hot outside.

This is an example of what the behavioral economist Daniel Kahneman called the "conjunction fallacy," an idea for which (among other ideas) he won the Nobel Prize. As soon as we hear the description of Linda, the idea of feminism pops into our minds, overwhelming our reason. It turns out that almost all of us are terrible reasoners about probability, in particular when we don't have the time to slow down and think. This can have terrible ramifications when we're taking a loan or negotiating a salary.

The second example is simpler. We human beings are, it turns out, deeply affected by the situations we're in and the other people around us, and often in surprising ways. Imagine that Dean Awn was a nefarious social scientist who wanted to use this gathering to replicate what's known as the "bystander effect." So Dean Awn decides to pump smoke into this room from under the door in order to see how long it would take for someone to actually do something about it—interrupt me, alert security, etc. If he did this, Dean Awn would find that it takes shockingly long for someone to speak up, so long as no one else is speaking up. If you were in a room by yourself, however, and the same smoke was pumped in, you'd react much faster, on average. There's no "good" reason for this. Rather, it's that we're the kind of creatures for whom being in certain situations causes us to act irrationally.

The last example is the most ominous in my estimation. It's the phenomenon of what's called implicit bias. It affects pretty much all of us, even if we're wholeheartedly committed to fairness and justice for all. Implicit biases are largely unconscious attitudes we have toward members of socially stigmatized groups, like black people, women, the elderly, and members of the LGBTQ community. Here's an example: if I asked everyone here to evaluate the same résumé, but the name on the résumé half of you received was "Mark Smith" and the name on the résumé the other half of you received was "Jamal Smith," one group's evaluations would be more positive than the others', and I'm sure you have no trouble guessing which is which.

Social stereotypes about race and gender and age and so on that are "out there" in the culture affect nearly all of us. They affect our feelings, our judgments, and our behavior. The evidence for this is pervasive and shocking. Doctors who score high

on tests of implicit bias (which is most doctors, because it's most everyone) are more likely to prescribe pain medications to white patients than black patients for equivalent symptoms and equivalent patient histories; police officers are more likely under time pressure to shoot unarmed black men than unarmed white men in a computer simulation; and on and on.

Please don't interpret me as saying that everyone is racist. Some forms of prejudice are the result of bad intentions, but not all. Some forms of prejudice are a result of the way we're wired up, that is, the way our brains work, in conjunction with the kind of world we've inherited, namely one with a long history of unjust social divisions.

What I *am* saying is that our minds are finite, that we're creatures—to use a fancy term—with “bounded rationality.” We're smart compared to some species, as Mahzarin Banaji puts it, but not smart enough when it comes to our own standards of rationality and morality. And this means, as Vitruvius said, that our instincts do indeed sometimes suck.

So where does this leave us? Can I give you any better advice than Vitruvius gave Emmett? I really can't, except to tell you—sorry to say it—that you're not special. *You* are in Emmett's bind, each of you, just like the rest of us. You're not magically exempt from suffering from the conjunction fallacy, the bystander effect, implicit bias, or the many other intellectual and moral frailties science has uncovered, just because you're a good guy or whatever. This means that you—and all of us—have got work to do. The work is figuring out when and how to trust our instincts, and how to cultivate them so that we may increase our self-trust.

Self-trust is the name of the game, I think, because it can take us to the heights of skill, to the passion and ability of the master builder. It's a dangerous game, though. I hope you will pursue it, but with real caution and humility.