

Free Will, or, Robert Sapolsky and His Bargain^[i]

Joseph Kugelmass



(Allan, 2016)^[ii]

“The definition of insanity,” a popular email signature used to tell us, “is doing the same thing and expecting a different result.” The quote was, invariably, attributed to Albert Einstein. Actually, it comes to us from a lesbian feminist who wrote mostly about her own life experiences. Her name was Rita Mae Brown, and her book, *Rubyfruit Jungle*, is a classic now. We get our most convincing modern definition of insanity from Brown.

Brown writes, “The trouble with Susan was that she made the same mistakes repeatedly.” She knows a quote about this, as provided, for her, by her fictional friend Jane Fulton. “Insanity is doing the same thing over and over again...” She just doesn’t let the quote stop her. Which is great *because* Jane Fulton is *right*.

Fulton’s theory holds up, almost. There is one exception to it: doing the same thing, over and over, in the presence of another human being. Then you get an interesting result. You get—*results*, plural. Part of being human is living with internal divides, including sudden ones. The “you” that sees one airplane overhead thinks nostalgically back on a

summer from your childhood. The “you” that sees four more planes, shortly thereafter, freaks out for half an hour.

If you go to someone’s house, and knock once, normally nobody will answer. (This is due to the breakdown of the social fabric, which started in 1913, for some reason.) If you knock twice, and the resident is home, you might get to the point where they open the door. Knock even once *more*, and you will either be treated like you are the police—sort of rolling the dice, that is—or else like you’re some dreaded relative everyone calls by their first name (e.g. “Aunt Darris”). And if it *is* the latter—a house full of people related to Aunt Darris—go on, knock all you like, but you will never get inside.

What you don’t have, past Level 1 of *The Sims*, is a series of robotic, hilarious re-dos where every time you knock, somebody opens that door, in the same genteel way, wearing the same vest and smile. (Heraclitus, a Greek, first noticed this and wrote about it in his seminal book, *This Is Not The Same River*.)

Timing and sequencing matter, because each iteration of even *an identical event* changes (for the other person) what they think is going on. Their behavior will also change. Ironically, therefore, you don’t introduce predictability into a system filled with living, thinking people, when you repeat a certain chain of events or stimuli over and over. That’s only a short-lived phase. Ultimately, what you get is a sudden, bifurcating shift, away from predictable responses, towards the unforeseeable—from two knocks, and a polite answer, to three knocks, when you suddenly begin careening into the fictional territory Susan colonizes, in Rita Mae Brown’s *Sudden Death*, after she stops taking Jane Fulton’s advice.

Why does this matter? Because it proves that even an identical stimulus does not, *contra* simplistic behaviorism, produce the same conditioned response every time, except under highly controlled conditions. If you put a dog in a kennel, and train it with a bell, you can get it to salivate whenever you ring the bell. If you put a dog in Minneapolis, and train it with a bell, it will follow The Salvation Army around, begging for table scraps, and (eventually) it will just stop—stop begging, stop salivating, its brain dissolving the underlying neurological arrangement that linked tinkling sounds with an incoming supper in the first place. That’s why, despite mentioning it frequently, Ivan Pavlov never actually visited Minneapolis.

This same kind of variance plagues *all* real-world problem solving. For example, to solve the problem of free will, Robert M. Sapolsky has published his new book *Determined: A Science of Life Without Free Will*.^[iii] The book moves at a Sapolsky-like pace,

blithely switching from longitudinal studies of behavioral triggers, which is how it opens, to a completely different series of chapters on emergent systems and chaos theory, just because Sapolsky is ready for something new. Something like chaos theory doesn't—I'm happy to report—spur Sapolsky to use, like he does near the end of his second chapter, that lovely noun phrase, "crack baby."

Sapolsky pities the crack baby. "Does society rush in to ensure that [you, a disadvantaged child] will be raised in relative affluence and with various therapies to overcome your neurodevelopmental problems? No" (42). Instead, your neurodevelopmental deficits get progressively worse.

As we know, your mother is likely to be drowning in the pathological consequences of her own miserable luck in life, with a good chance of leaving you neglected, abused, shuttled through foster homes. Well, does society at least mobilize then to counterbalance that additional bad luck, ensuring that you live in a safe neighborhood with excellent schools? Nope, your neighborhood is likely to be gang-riddled and your school underfunded. (43)

I'm not going to try attacking this logic on the grounds that it is, simultaneously, horrifically condescending to the poor, *and* sort of indifferent to the very resolvable social ills that Sapolsky is so eager to list out. At least he *does* list a few of them. The rest go off and drench some insolvent, waterlogged mother. I can see her now, leaning tiredly against one of his cheap metaphors, drowning in her "own miserable luck."

The result of all this unsolved social unfairness is that people who grow up in traumatic, abusive, and stressful circumstances have—according to Sapolsky, once he finally comes back around to them—poor impulse control.

Developing post-traumatic stress disorder after trauma transforms the amygdala. Synapse number increases along with the extent of the circuitry by which the amygdala influences the rest of the brain. The overall size of the amygdala increases, and it becomes more excitable, with a lower threshold for triggering fear, anxiety, and aggression. (58)

Sure. I've seen a thousand transformed amygdalas, and I've diagnosed them all. But what if you just have a really lousy childhood?

Lots of childhood stress, by way of glucocorticoids, impairs construction of the frontal cortex, producing an adult less adept at helpful things like impulse control. Lots of exposure to testosterone early in life makes for the construction of a highly reactive amygdala, producing an adult more likely to respond aggressively to provocation. (65)

All Sapolsky is really telling us, here, is that if you look closely at an individual's brain, you can sometimes tell whether or not they've learned to live more according to their nerves—like someone trying, right down to their neurons, to guard themselves against some fresh hell of trauma or hunger—or more according to their own pleasant rules for a well-ordered life. This is not news, and it's not even true, as Sapolsky would have us believe, that a sensitized amygdala (for example) is a sign of neurological disease. It's just less common than other configurations found in other people's brains. It's the brain's response to real circumstances out there, in the world. If the individual's environment shifts, and they're suddenly awash in comfort, then it's certainly possible that they'll mourn a certain *tendency* in themselves towards poor impulse control. But not necessarily, obviously. Not everyone has to put impulse control higher on the scale of evolved cognition than, for instance, the ability to compose music.

More importantly, such a tendency is just one mind's ready-to-go, already patterned reaction to past events. It's not inevitable, and the brain doesn't make it so, as anyone looking at any relevant study can see (because they deal, not in absolutes, but in probable outcomes). The difference between past and present, comprehended and encoded within our brains, is the difference between *reduced impulse control* (in an environment where it may be, or at least *have once been*, actually *disadvantageous*) and *zero impulse control*. That literally can only be simulated in a laboratory, by doing things to people's brains (like using chemicals to power sections of one's frontal lobe down) that almost *never* happen in the normal course of a day, except at certain matinees.

In any event, for Sapolsky, the whole question of volition seems to hinge on the brain's capacity under particular circumstances to hold one part of itself (e.g. the amygdala) in check with another part (e.g. the pre-frontal cortex). I guess his theory is that if you can't stop yourself from doing *something*, especially something *bad*, then all you are really lacking is an opportunity—no matter what time of year or day it is, and no matter where you are.

In fact, though, there are a hundred reasons why someone with poor impulse control (a burned out social worker, say, who has just quit her job) will suddenly land in the middle of a neighborhood that values restraint (a gated residential community, where a friend lives). At that point, the "disorderly" visitor become an outlier, quite unlike the otherwise predictable folks around them. Inside the gates everyone goes, at six p.m., to the gazebo. There's free coffee and crumb cake; meanwhile, the newly unemployed visitor is driving herself and her friend to Baja on the strength of her severance check.

In the walk-in clinic, which she left, everyone is (comparatively) stressed. When a new person arrives, the front doorbell goes off. In the waiting room, people look up like they've just been handed a bomb threat. One man, Shaun, does not. He is chill.

So, then, who is "more" free here: the condo crowd, having coffee, or the crowd of intake clients, in a waiting room, shuddering over the awful door alarm? Who is the greater outlier and the less predictable exception—Shaun, who is chill, or the ex-social worker, who is going to wake up hungover on a beach? Who is exercising the maximum degree of freedom, given the situation?

I should mention, before you really start work on those rhetorical questions, that it's not necessary to ask Shaun, or the former social worker, or anyone holding a piece of crumb cake, whether they *feel* free, or whether they *perceive themselves* to be a responsible, autonomous agent. That shouldn't make any difference; a person, presumably, might also feel extremely purple, without being correct about themselves in that regard. So you can safely toss aside each and every new version of a very old study, in which people argue for a feeling of freedom, and then do something predictable. According to *every* study, ever performed, people *are* sometimes wrong. (Sapolsky, however, devotes an entire chapter to documenting, and citing, this amazing tendency in people.)

Anyhow, back to the question of actual freedom, as opposed to something merely perceived and then self-reported. Is this freedom manifest? Is it there in differing amounts? One way of answering this is by sharply distinguishing between the absolute condition of *the freedom to act*, described admirably by philosophy (Existentialism, in particular) – and the tendencies and appetites of daily life. Daily life has its routines, its endearing room for preferences; it is not a waste of freedom to participate in it, but on a dull day, one's capacity for freedom is not necessarily visibly or sorely tested. That doesn't change the nature of freedom or our ability to possess it. Driving a high-performance sports car half a mile, per day, to a Park 'n Go, does not demonstrate that you are, in fact, driving to catch your light-rail in a golf cart.

The other way is to ask three questions: first of all, what factor (we'll call it, "Factor X") is the biggest predictor of a certain behavioral trend? Second, how much influence does Factor X really have? And, finally, how stable is Factor X, both literally, and in the minds of people nearby?

In other words, look at freedom from above. Don't try, like Sapolsky, to disprove it by asking people to go chasing (mentally) after fleeting, nebulous *sensations* of agency, and then doing an MRI scan to see what little node in their pre-frontal cortex *gives* them

these fleeting sensations of agency. Instead, try to predict *what they will do*, being honest if you're wrong, especially if you're wrong more than 40% of the time, like the other scientists who end up doing most of Sapolsky's legwork for him.

You will, in every case, get the same results. First of all, over time, as your data set grows, you will get more *kinds* of outliers -- even if the degree of difference, and the percentage of actors you predict correctly, stays perfectly steady. That's a problem because you aren't tracking those causal links, and at some point, if they re-intersect with your big prediction, you won't be ready.

But there's another problem, too. As I mentioned, earlier, the same stimulus can trigger different, even contrasting or opposite associations. Environmental triggers become part of an invisible narrative; there is basically no way to predict how future iterations of the same environment will affect *the same exact people*. At some point, if you just keep on betting they'll do more of the same, you'll be proved totally and absolutely wrong.

If it's nighttime, and there's one place to buy hamburgers, then at least half of the drivers traveling at night will stop. That can tell you approximately how many hamburgers to buy. But it can't trace the lines that start where each car's headlights are blazing, and continue out from there, towards the horizon, one dinner in solitude, another on the lam.^[iv] As Rita Mae Brown once put it, memorably, "you may need to be lumped in a group politically ... I understand that, but I don't accept it." You don't know why some people stop, and others don't. For now, perhaps, that doesn't matter. Tomorrow the whole stand will up and vanish. You'll drive by a rusting old diesel pump, and that'll be the full extent of it.

What I find so strange, and sad, about Robert Sapolsky's new book is that all *he* is trying to do, by writing this, is to free himself from the supposition that everyone faces equal opportunities in life. Of course they don't. "I haven't believed in free will since adolescence," he writes, like a certain kind of published vegetarians, "and it's been a moral imperative for me to view humans without judgment or the belief that anyone deserves anything special, to live without a capacity for hatred or entitlement" (9). I'm not sure why Sapolsky's moral imperative requires him to explain the nature of the amygdala, however, while ignoring (for example) the function of memory in the creation of new perceptions. Nor am I certain that one must, to live "without a capacity for hatred or entitlement," go forth and doggedly pursue the argument that one was right as a teenager, *is still right*, and can prove it with a mountain of identically meaningful, and irrelevant, studies copped from Big Data. But off Sapolsky goes. Everyone from chaos theorists to

quantum physicists just don't understand what it is...*to choose*. It's impossible, actually. Why can't they just *see this*, as clearly as he does?

All I know for sure is that it is *not* a moral imperative for Robert Sapolsky to achieve this perception of compassionate equivalence by paying with his freedom. He is free to exercise as much, or as little, moral compassion as he wants, at all times, no matter how old he is. That's the problem, I suppose. To submerge oneself in the unthinkable complexity of a world inhabited by more than 7.5 billion free actors. To imagine all these human beings as equals, *without* basing all that on some trumped-up lack, in our world that is panting from other, realer insufficiencies. To see how little, for people without his stratospheric concerns, *their* existential freedom really entitles them to buy, or how laughable they might find Sapolsky's bargain, even in a seller's market. To most people, even teenagers, what Sapolsky has attempted, merely *attempted*, to do, is the very definition of insanity.

NOTES

[i] An earlier version of this essay was also published in *Splice Today*, available online at URL = <<https://www.splicetoday.com/writing/free-will-robert-sapolsky-and-his-bargain>>.

[ii] Leslie Allan, "Psychological Research on Free Will Intuitions: A Critical Review," *Rational Realm*, available online at URL = <<https://www.rationalrealm.com/philosophy/metaphysics/psychological-research-free-will-intuitions.html>>.

[iii] R. Sapolsky, *Determined: A Science of Life Without Free Will* (London: Penguin, 2023).
Pagination for quoted texts cite this edition.

[iv] See E. Hopper, *Nighthawks*, for more on this.