One

The Behavior

We have our strategy in place. A behavior has occurred—one that is reprehensible, or wonderful, or floating ambiguously in between. What occurred in the prior second that triggered the behavior? This is the province of the nervous system. What occurred in the prior seconds to minutes that triggered the nervous system to produce that behavior? This is the world of sensory stimuli, much of it sensed unconsciously. What occurred in the prior hours to days to change the sensitivity of the nervous system to such stimuli? Acute actions of hormones. And so on, all the way back to the evolutionary pressures played out over the prior millions of years that started the ball rolling.

So we're set. Except that when approaching this big sprawling mess of a subject, it is kind of incumbent upon you to first define your terms. Which is an unwelcome prospect.

Here are some words of central importance to this book: aggression, violence, compassion, empathy, sympathy, competition, cooperation, altruism, envy, schadenfreude, spite, forgiveness, reconciliation, revenge, reciprocity, and (why not?) love. Flinging us into definitional quagmires.

Why the difficulty? As emphasized in the introduction, one reason is that so many of these terms are the subject of ideological battles over the appropriation and distortions of their meanings.[*1] Words pack power and these definitions are laden with values, often wildly idiosyncratic ones. Here's an example, namely the ways I think about the word "competition": (a) "competition"—your lab team races the Cambridge group to a discovery (exhilarating but embarrassing to admit to); (b) "competition"—playing pickup soccer (fine, as long as the best player shifts sides if the score becomes lopsided); (c) "competition"—your child's teacher announces a prize for the best outlining-your-fingers

Thanksgiving turkey drawing (silly and perhaps a red flag—if it keeps happening, maybe complain to the principal); (d) "competition"—whose deity is more worth killing for? (try to avoid).

But the biggest reason for the definitional challenge was emphasized in the introduction—these terms mean different things to scientists living inside different disciplines. Is "aggression" about thought, emotion, or something done with muscles? Is "altruism" something that can be studied mathematically in various species, including bacteria, or are we discussing moral development in kids? And implicit in these different perspectives, disciplines have differing tendencies toward lumping and splitting—these scientists believe that behavior X consists of two different subtypes, whereas those scientists think it comes in seventeen flavors.

Let's examine this with respect to different types of "aggression."[2] Animal behaviorists dichotomize between offensive and defensive aggression, distinguishing between, say, the intruder and the resident of a territory; the biology underlying these two versions differs. Such scientists also distinguish between conspecific aggression (between members of the same species) and fighting off a predator. Meanwhile, criminologists distinguish between impulsive and premeditated aggression. Anthropologists care about differing levels of organization underlying aggression, distinguishing among warfare, clan vendettas, and homicide.

Moreover, various disciplines distinguish between aggression that occurs reactively (in response to provocation) and spontaneous aggression, as well as between hot-blooded, emotional aggression and cold-blooded, instrumental aggression (e.g., "I want your spot to build my nest, so scram or I'll peck your eyes out; this isn't personal, though").[3] Then there's another version of "This isn't personal"—targeting someone just because they're weak and you're frustrated, stressed, or pained and need to displace some aggression. Such third-party aggression is ubiquitous—shock a rat and it's likely to bite the smaller guy nearby; a beta-ranking male baboon loses a fight to the alpha, and he chases the omega male;[*] when unemployment rises, so do rates of domestic violence. Depressingly, as will be discussed in chapter 4, displacement aggression can decrease the perpetrator's stress hormone levels; giving ulcers can help you avoid getting them. And of course there is the ghastly world of aggression that is neither reactive nor instrumental but is done for pleasure.

Then there are specialized subtypes of aggression—maternal aggression, which often has a distinctive endocrinology. There's the difference between

aggression and ritualistic *threats* of aggression. For example, many primates have lower rates of actual aggression than of ritualized threats (such as displaying their canines). Similarly, aggression in Siamese fighting fish is mostly ritualistic.*

Getting a definitional handle on the more positive terms isn't easy either. There's empathy versus sympathy, reconciliation versus forgiveness, and altruism versus "pathological altruism."[4] For a psychologist the last term might describe the empathic codependency of enabling a partner's drug use. For a neuroscientist it describes a consequence of a type of damage to the frontal cortex—in economic games of shifting strategies, individuals with such damage fail to switch to less altruistic play when being repeatedly stabbed in the back by the other player, despite being able to verbalize the other player's strategy.

When it comes to the more positive behaviors, the most pervasive issue is one that ultimately transcends semantics—does pure altruism actually exist? Can you ever separate doing good from the expectation of reciprocity, public acclaim, self-esteem, or the promise of paradise?

This plays out in a fascinating realm, as reported in Larissa MacFarquhar's 2009 *New Yorker* piece "The Kindest Cut."[5] It concerns people who donate organs not to family members or close friends but to strangers. An act of seemingly pure altruism. But these Samaritans unnerve everyone, sowing suspicion and skepticism. Is she expecting to get paid secretly for her kidney? Is she that desperate for attention? Will she work her way into the recipient's life and do a *Fatal Attraction*? What's her deal? The piece suggests that these profound acts of goodness unnerve because of their detached, affectless nature.

This speaks to an important point that runs through the book. As noted, we distinguish between hot-blooded and cold-blooded violence. We understand the former more, can see mitigating factors in it—consider the grieving, raging man who kills the killer of his child. And conversely, affectless violence seems horrifying and incomprehensible; this is the sociopathic contract killer, the Hannibal Lecter who kills without his heart rate nudging up a beat.*[6] It's why *cold-blooded* killing is a damning descriptor.

Similarly, we expect that our best, most prosocial acts be warmhearted, filled with positive affect. Cold-blooded goodness seems oxymoronic, is unsettling. I was once at a conference of neuroscientists and all-star Buddhist monk meditators, the former studying what the brains of the latter did during meditation. One scientist asked one of the monks whether he ever stops meditating because his knees hurt from all that cross-leggedness. He answered,

“Sometimes I’ll stop sooner than I planned, but not because it hurts; it’s not something I notice. It’s as an act of kindness to my knees.” “Whoa,” I thought, “these guys are from another planet.” A cool, commendable one, but another planet nonetheless. Crimes of passion and good acts of passion make the most sense to us (nevertheless, as we shall see, dispassionate kindness often has much to recommend it).

Hot-blooded badness, warmhearted goodness, and the unnerving incongruity of the cold-blooded versions raise a key point, encapsulated in a quote from Elie Wiesel, the Nobel Peace Prize winner and concentration camp survivor: “The opposite of love is not hate; its opposite is indifference.” The biologies of strong love and strong hate are similar in many ways, as we’ll see.

Which reminds us that we don’t hate aggression; we hate the wrong kind of aggression but love it in the right context. And conversely, in the wrong context our most laudable behaviors are anything but. The motoric features of our behaviors are less important and challenging to understand than the meaning behind our muscles’ actions.

This is shown in a subtle study.[7] Subjects in a brain scanner entered a virtual room where they encountered either an injured person in need of help or a menacing extraterrestrial; subjects could either bandage or shoot the individual. Pulling a trigger and applying a bandage are different behaviors. But they are similar, insofar as bandaging the injured person and shooting the alien are both the “right” things. And contemplating those two different versions of doing the right thing activated the same circuitry in the most context-savvy part of the brain, the prefrontal cortex.

And thus those key terms that anchor this book are most difficult to define because of their profound context dependency. I will therefore group them in a way that reflects this. I won’t frame the behaviors to come as either pro- or antisocial—too cold-blooded for my expository tastes. Nor will they be labeled as “good” and “evil”—too hot-blooded and frothy. Instead, as our convenient shorthand for concepts that truly defy brevity, this book is about the biology of our best and worst behaviors.